This text is intended to introduce the reader to a wide variety of credibility models and, in so doing, trace the historical development of the subject. The Bayesian approach to statistics is emphasized, revealing the author’s personal preference. The reader should be able to use this work as a foundation for understanding more sophisticated treatments in other works. For example, by seeing how various formulas are derived in the Bayesian paradigm, the reader should be able to understand other works describing the Bayesian approach to credibility. Another goal is to present the key assumptions underlying the various credibility models and to discuss the advantages and disadvantages of the various approaches.

This work is intended to be largely self-contained. Although numerous references to the technical literature are provided, few are necessary for an understanding of the material discussed here. Rather, they are provided for those who would like to consult original sources and/or obtain some insight into the more advanced topics omitted from this introductory work. A large number of exercises are provided to help reinforce understanding of the material. Most of these have been taken from past examinations of the Casualty Actuarial Society. Complete solutions to all of the text exercises are available in a companion solutions manual.

The emphasis in the first ten chapters of this introductory text is on basic statistical concepts.

In Chapter 1, we (a) discuss two major statistical paradigms, (b) offer a glimpse into the nature of credibility, (c) introduce a simple practical problem later solved in Chapter 6 using credibility procedures, and (d) present a brief review of the key historical developments in credibility theory and its application to practical insurance problems.
In Chapter 2, we review the basic concepts of Bayesian analysis, and in Chapter 3 we discuss statistical loss functions. In Chapter 4, we use an example originally employed by Hewitt [1970] to illustrate the use of Bayesian concepts in the insurance ratemaking process. The key ideas are the use of the predictive distribution of aggregate claim amounts and the use of the (Bayesian) conditional mean to estimate pure premium amounts.

In Chapter 5, we describe the limited fluctuation credibility model. This is primarily of historical interest, because it is not in wide use today.

In Chapter 6, we present the development of an alternative credibility model proposed by Bühlmann [1967] as well as a special case of a more general model proposed by Bühlmann and Straub [1972]. The general Bühlmann-Straub model is presented in Chapter 7.

In Chapter 8, we discuss an important general result of Ericson [1970]. It turns out that some specific results described in Mayerson [1964] are simply special cases of Ericson’s more general result. We also present an example which shows that the Bühlmann estimate does not always equal the corresponding Bayesian estimate. In Chapter 9, we use the statistical machinery developed to describe Ericson’s result to construct the predictive distribution of a more realistic two-stage model. Here the number of claims is assumed to follow a Poisson distribution and the claim amounts are based on an exponential distribution.

In Chapter 10, we show that Bühlmann’s model produces least squares linear approximations to the Bayesian estimate of the pure premium.

In the first ten chapters, we do not discuss important practical issues such as how to apply these procedures in dealing with issues likely to be encountered in real-life situations. Moreover, we do not attempt to discuss more sophisticated theoretical concepts such as multivariate extensions of the results presented here. These are all left for a more advanced treatment in later chapters and elsewhere.

The three prior editions of this text have been in use since 1994. Our goal over all of these years has been to make this text of practical use to the working actuary/actuarial student. To further this goal, we have added three new chapters to this fourth edition. Each of Chapters 11 through 15 now deals in depth with a practical application of the
concepts developed earlier in the text. Chapter 11 discusses a Bayesian procedure for comparing two binomial proportions. Many researchers, including us, feel that this is far superior to the frequentist scheme of deciding whether or not to reject the null hypothesis that the two proportions are equal. Chapter 12 describes a procedure suggested by Fuhrer [1988] that has application to health insurance. Chapter 13 summarizes work completed by Rosenberg and Farrell [2008] that describes a scheme for predicting the frequency and severity of hospitalization cost for a group of young children suffering from cystic fibrosis. Chapters 14 and 15 continue from earlier editions. In Chapter 14, we use the concept of conjugate prior distributions to estimate probabilities arising from a data quality problem. In Chapter 15, we present an application of empirical Bayesian procedures to a problem in automobile insurance ratemaking. This is the application and solution proposed by Morris and van Slyke [1979].

The other major change in the fourth edition is in Chapter 1 where we have expanded the historical discussion surrounding the work of Bayes himself as well as Laplace.

We assume here that the reader has a working knowledge of (1) the integration techniques normally taught during the second semester of a university course in calculus, (2) basic probability and statistics as taught during a two-semester university course having a calculus prerequisite, and (3) matrix manipulation and multiplication techniques. In particular regard to integral calculus, we assume that the reader can perform integration by parts and integration by change of variable (as done in Section 9.4). We also note that in a number places we have changed either the order of integration or summation. In general such manipulation requires the verification of one or more conditions to ensure that the sums actually converge. Because we are dealing here with probabilities that sum to one, we are never in danger of diverging to infinity. Hence, we will omit the verification step in this work.

For many years, the topic of credibility theory has been included in the preliminary exams jointly sponsored by the Casualty Actuarial Society and the Society of Actuaries. All of the topics in that collection as well as the general learning objectives and the sample exam questions are thoroughly covered by this text. It has thus been an officially approved reference for this topic for most of those years.
Many people have contributed to the completion of this project. In particular I am grateful to the late Professor James C. Hickman, FSA, ACAS, University of Wisconsin, and Gary G. Venter, FCAS, ASA, CERA, MAAA for their generous assistance with an initial version of this work. Jim taught me a lot about Bayesian statistics and Gary introduced me to the Bühlmann and limited fluctuation approaches. Gary also read this and earlier versions and suggested a number of the figures that appear in Chapters 6 and 10.

I also want to thank Professor Donald B. Rubin of Harvard University for introducing me to the Bayesian paradigm. One of my other early teachers was Professor Dennis V. Lindley, at George Washington University. I would also like to express my appreciation to Professor James C. Owings, Jr., my dissertation advisor in the Department of Mathematics at the University of Maryland.

In addition to Gary Venter, I am also indebted to the other members of the initial editorial review team for their suggestions that improved the exposition and notation of this work. This group included Professor Elias Shiu, ASA, University of Iowa, Professor Samuel H. Cox, FSA, University of Manitoba, Professor Larry Santoro, College of Insurance, Professor Michael R. Powers, Temple University, and David F. Mohrman, FCAS, and Stephen W. Philbrick, FCAS, both of Towers Watson. Professor Shiu supplied me with the clever proofs presented in Sections 3.2 and 5.2.

The fourth edition was also reviewed by Steven Craighead, Margie Rosenberg, Dennis Tolley, Chuck Fuhrer, Bob Cumming, Joan Barrett, and Vincent Kane. The author is grateful for their generous and thoughtful assistance with this text.

Finally, I would like to thank the staff of ACTEX Publications, Inc, especially Marilyn Baleshiski and Gail Hall for their assistance with the publication of this text.

Reston, Virginia  
Thomas N. Herzog, Ph.D., ASA  
June 2010
CHAPTER 3  LOSS FUNCTIONS  31

3.1 Squared Error Loss Functions ....................................................... 31
3.2 Absolute Error Loss Functions ...................................................... 34
3.3 Almost Constant Loss Functions .................................................. 36
3.4 Other Loss Functions ................................................................. 38
3.5 Further Remarks ............................................................................ 39
3.6 Exercises ..................................................................................... 40

CHAPTER 4  DISCRETE FREQUENCY-SEVERITY INSURANCE MODEL UNDER INDEPENDENCE  43

4.1 Initial Pure Premium Estimates ..................................................... 44
4.2 Revised Pure Premium Estimates and Predictive Distributions .... 46
4.3 Additional Remarks ................................................................. 50
4.4 Exercises ..................................................................................... 51

CHAPTER 5  LIMITED FLUCTUATION CREDIBILITY APPROACH  53

5.1 Introduction ................................................................................ 53
5.2 Mathematical Preliminaries .......................................................... 53
5.3 The Full Credibility Approach ...................................................... 56
  5.3.1 The Rationale for Full Credibility........................................ 56
  5.3.2 The Full Credibility Formula ............................................. 57
  5.3.3 Numerical Examples ........................................................ 59
  5.3.4 Minimum Number of Expected Claims Required for Full Credibility .............................................. 61
  5.3.5 Summary of the Full Credibility Approach ....................... 62
5.4 The Partial Credibility Approach .................................................. 62
  5.4.1 The Rationale for Partial Credibility ................................. 62
  5.4.2 The Partial Credibility Formula ....................................... 62
  5.4.3 Partial Credibility Example .............................................. 64
  5.4.4 Summary of the Partial Credibility Approach .................. 64
5.5 Further Remarks and Summary .................................................. 65
5.6 Exercises ..................................................................................... 66
CHAPTER 6  BÜHLMANN’S APPROACH  71

6.1 Introduction ................................................................................... 71
6.2 Basic Elements of Bühlmann’s Approach ........................................ 71
   6.2.1 Bühlmann’s Credibility Formula ..................................... 71
   6.2.2 Hypothetical Means ......................................................... 72
   6.2.3 Variance of the Hypothetical Means ................................ 74
   6.2.4 Process Variance .............................................................. 78
   6.2.5 Expected Value of the Process Variance ......................... 78
   6.2.6 Credibility Estimates........................................................ 82
   6.2.7 Summary of the Bühlmann Approach .............................. 84
6.3 Characteristics of the Credibility Factor ........................................ 85
6.4 Further Examples of Bühlmann Estimates .................................... 85
6.5 Components of the Total Variance ............................................... 89
6.6 Practical Applications of Bühlmann’s Model ............................... 90
   6.6.1 General Approach ............................................................ 91
   6.6.2 Summary of the General Approach.................................. 95
   6.6.3 Using Widely-Available Software to Compute Bühlmann Estimates ............................................................ 96
   6.6.4 Ad Hoc Methods .............................................................. 97
6.7 Two Equivalent Ways to Calculate Z............................................ 99
6.8 Further Comments ....................................................................... 102
6.9 Exercises ..................................................................................... 104

CHAPTER 7  BÜHLMANN-STRAUB MODEL  113

7.1 Introduction ................................................................................... 113
7.2 Model for One Policyholder with Known Underlying Probability Distributions ........................................ 113
   7.2.1 Basic Elements ............................................................... 113
   7.2.2 Examples ........................................................................ 115
7.3 Model for Two or More Policyholders ....................................... 118
   7.3.1 Basic Elements ............................................................... 118
   7.3.2 Examples ........................................................................ 121
   7.3.3 Summary ........................................................................ 128
   7.3.4 What if the Estimate of the Variance Is Negative? .... 129
7.4 Semiparametric Estimation: Simplifying the Estimation of the Parameters of the Model of Section 7.3 .... 130
7.5 Exercises ..................................................................................... 133
CHAPTER 8  CREDIBILITY AND BAYESIAN INFERENCE  137

8.1 Introduction ................................................................................. 137
8.2 Conjugate Prior Distributions ..................................................... 137
  8.2.1 Bernoulli Example ............................................................ 139
  8.2.2 Poisson Example ............................................................. 140
  8.2.3 Normal Example .............................................................. 142
8.3 Credibility and Conjugate Distributions ..................................... 144
  8.3.1 Bernoulli Example Revisited ............................................ 145
  8.3.2 Poisson Example Revisited .............................................. 147
  8.3.3 Normal Example Revisited .............................................. 148
8.4 Credibility and Exponential Families .......................................... 149
  8.4.1 Exponential Families........................................................... 149
  8.4.2 Conditions for the Equality of the
       Bühlmann and Bayesian Estimators.............................. 151
  8.4.3 An Alternative Procedure for Obtaining
       Bühlmann and Bayesian Estimates .............................. 152
8.5 Summary of Key Results of the Bayesian Paradigm .................... 152
  8.5.1 Parameters ..................................................................... 152
  8.5.2 Random Variables ................................................................ 152
  8.5.3 Pure Premium Estimates (ExpectedValues)........................ 153
  8.5.4 Example 6.1 Revisited ................................................... 153
8.6 Exercises ..................................................................................... 155

CHAPTER 9  FREQUENCY-SEVERITY INSURANCE MODEL
            WITH CONTINUOUS SEVERITY COMPONENT  163

9.1 Introduction ................................................................................. 163
9.2 Predictive Distributions ............................................................... 163
9.3 Frequency Component ................................................................ 164
9.4 Severity Component .................................................................. 166
9.5 The Two-Stage Model ................................................................ 168
9.6 Pure Premium Estimator ............................................................. 170
9.7 Alternative Pure Premium Estimator
       Using Bühlmann’s Model ................................................ 170
9.8 Summary of Chapters 8 and 9 ..................................................... 173
9.9 Advantages and Disadvantages of the
       Bühlmann and Bayesian Approaches ............................. 174
9.10 Bühlmann View ......................................................................... 175
9.11 Exercises ..................................................................................... 176
CHAPTER 10 CREDIBILITY AND LEAST SQUARES 183
10.1 Introduction.......................................................................................... 183
10.2 Linear Approximation Example.......................................................... 183
10.3 Linear Approximation Theory .......................................................... 185
10.4 Exercises ......................................................................................... 190

CHAPTER 11 CASE STUDY: CLINICAL TRIALS – AN APPLICATION OF BAYESIAN INFERENCE ABOUT BINOMIAL PROPORTIONS 193
11.1 Introduction.......................................................................................... 193
11.2 Background ....................................................................................... 195
11.3 Harvard Study Data and Analysis ...................................................... 196
11.4 Analysis Using a Uniform Prior Distribution .................................... 196
11.5 Analysis Using a More-Informative Prior Distribution ..................... 200
11.6 Conclusions and General Observations ........................................... 202
11.7 Exercises ......................................................................................... 203

CHAPTER 12 CASE STUDY: APPLICATION OF CREDIBILITY TO GROUP HEALTH INSURANCE 205
12.1 Introduction.......................................................................................... 205
12.2 General Approach ........................................................................... 206
  12.2.1 General Approach when \( n = 1 \) ................................................. 207
  12.2.2 General Approach when \( n > 1 \) ............................................... 208
12.3 Credibility for Group Insurance ....................................................... 210
  12.3.1 Notation ................................................................................. 210
  12.3.2 Credibility Model for \( E[X_2 | \bar{X}_1 = \bar{x}_1] \) ............................ 210
12.4 Specific Approach to Obtaining an Optimal Pooling Level .......... 213
12.5 A More General Model ................................................................... 217
12.6 Summary and Conclusion ............................................................... 218
12.7 Exercises ......................................................................................... 218
## Chapter 13: A Case Study in Bayesian Analysis:
The Impact of Newborn Screening on Cystic Fibrosis Hospitalization Costs

**13.1 Introduction**

- **13.1.1 Purpose**
- **13.1.2 Data**
- **13.1.3 The Model**
- **13.1.4 Method Used to Estimate Model Parameters**
- **13.1.5 WinBUGS Software Package**
- **13.1.6 Frequency Stage**
- **13.1.7 Severity Stage**
- **13.1.8 Predictor Variables**
- **13.1.9 Regression Model for Frequency Component**
- **13.1.10 Regression Model for Physician Costs**
- **13.1.12 Incorporating the Variability of the Estimates of the Parameters of the Regression Model**

**13.2 Results**

- **13.2.1 Predictive Distribution for the Number of Hospital Visits of Child B During its First Year of Life**
- **13.2.2 Predictive Distribution for the Cost of Each Hospitalization of Child B During its First Year of Life**
- **13.2.3 Predictive Distribution for the Cost of Hospitalization of Child B During its First Year of Life**

**13.3 Markov Chain Monte Carlo Techniques**

## Chapter 14: A Case Study:
The Miscoded Mortgage Records

**14.1 Introduction**

- **14.1.1 Background**
- **14.2.1 Statement of the Problem**
- **14.1.3 Overview of the Methodological Approach**

**14.2 Beta Prior Densities and Conditional Probabilities**

**14.3 The Uniform Density Function**

**14.4 Concluding Remarks on Case Study**
14.5 A Similar Application: Loss Ratios ............................................ 241
  14.5.1 Notation .......................................................................... 241
  14.5.2 Assumptions ................................................................... 242
  14.5.3 Posterior Probability ...................................................... 242
  14.5.4 Frequentist Approach to this Problem............................ 243
  14.5.5 Alternative Prior Distribution ........................................ 243
14.6 Exercises ..................................................................................... 245

CHAPTER 15 MORRIS - VAN SLYKE ESTIMATION  247

15.1 Introduction ................................................................................. 247
15.2 The Morris - van Slyke Model .................................................... 248
  15.2.1 Iterative Procedure ......................................................... 251
  15.2.2 Final Estimators of the Credibility Factors .................... 253
  15.2.3 Final Estimators of the \(\alpha_i\)’s ........................................... 253
  15.2.4 Measures of Fit............................................................... 254
15.3 Worked Solution of Automobile Rating Class Example ............ 254
15.4 Summary of the Morris - van Slyke Model ...................... 256
15.5 Further Remarks .......................................................................... 257
15.6 Exercises ..................................................................................... 257

APPENDIX A: PREDICTIVE MODELING OF COSTS
FOR A CHRONIC DISEASE
WITH ACUTE HIGH-COST EPISODES  261

Abstract ................................................................................................. 261
A.1 Introduction....................................................................................... 262
A.2 Methods........................................................................................... 263
  A.2.1 Data ................................................................................ 263
  A.2.2 Model ............................................................................. 267
  A.2.3 Prediction ....................................................................... 269
A.3 Results ........................................................................................ 271
A.4 Conclusion .................................................................................. 287
A.5 Acknowledgments ....................................................................... 288
Appendix: A Markov Chain Monte Carlo Technique .................. 289
References ........................................................................................... 291
APPENDIX B: COMPARING CREDIBILITY ESTIMATES OF HEALTH INSURANCE CLAIMS COSTS 293

Abstract .................................................................................................................. 293
B.1 Background ..................................................................................................... 294
B.2 Introduction ..................................................................................................... 294
B.3 The Data .......................................................................................................... 295
B.4 The Models ...................................................................................................... 296
   B.4.1 The Linear Mixed Model ........................................................................... 296
   B.4.2 The Hierarchical Bayes Model ................................................................. 302
B.5 Results .............................................................................................................. 306
   B.5.1 The Groups ............................................................................................. 306
   B.5.2 The Chain ............................................................................................... 307
   B.5.3 The Parameters ....................................................................................... 307
   B.5.4 Predicted Values ...................................................................................... 309
B.6 Discussion ......................................................................................................... 313
Acknowledgments ................................................................................................... 314
References ............................................................................................................... 314

ANSWERS TO THE EXERCISES ......................................................................... 317
REFERENCES ......................................................................................................... 327
INDEX ...................................................................................................................... 335
CHAPTER 1

INTRODUCTION AND HISTORY

1.1 INTRODUCTION

According to Rodermund [1989, page 3], “the concept of credibility has been the casualty actuaries’ most important and enduring contribution to casualty actuarial science.”

In order to present a brief history of credibility, it will be helpful to begin by describing two major statistical paradigms and three major approaches to credibility. This will facilitate our description of the historical development.

1.2 STATISTICAL PARADIGMS

Credibility is an example of a statistical estimate. Statistical estimates are obtained through the use of statistical formulas or models which, in turn, are based on statistical approaches or paradigms. There are two major statistical paradigms of current interest, which are (a) the frequentist or classical paradigm, and (b) the Bayesian paradigm.

In the frequentist paradigm, the probability of an event is based on its relative frequency. All prior and/or collateral information is ignored. Proponents of the frequentist paradigm view it as being objective, because all attention is devoted to the observations (data). Some of the key constructs of the frequentist paradigm are the Neyman-Pearson Lemma, tests of statistical hypotheses, confidence intervals, and unbiased estimates.

In the Bayesian paradigm, probability is treated as a rational measure of belief. Thus, the Bayesian paradigm is based on personal or subjective probabilities and involves the use of Bayes’ theorem. Prior and/or collateral information is incorporated explicitly into the model via the prior distribu-
tion and the likelihood. Some of the key constructs of the Bayesian paradigm, in addition to Bayes’ theorem itself, are conditional probabilities, prior distributions, predictive distributions, and (posterior) odds ratios.

1.3 WHAT IS CREDIBILITY?

Suppose we have two collections of data, as illustrated in the following figure.

<table>
<thead>
<tr>
<th>Prior Observations</th>
<th>Current Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>#  #  #  #</td>
<td>#  #  #  #</td>
</tr>
<tr>
<td>#  #  #  #</td>
<td>#  #  #  #</td>
</tr>
<tr>
<td>#  #  #  #</td>
<td>#  #  #  #</td>
</tr>
<tr>
<td>#  #  #  #</td>
<td>#  #  #  #</td>
</tr>
</tbody>
</table>

**FIGURE 1.1**

One collection consists of current observations, taken from the most recent period of observation. The second collection has observations for one or more prior periods. The various approaches to credibility give us different “recipes” for combining the two collections of observations to obtain an overall estimate.

Under some approaches to credibility, a compromise estimator, \( C \), is calculated from the relationship

\[
C = ZR + (1-Z)H, \tag{1.1}
\]

where \( R \) is the mean of the current observations (for example, the data), \( H \) is the prior mean (for example, an estimate based on the actuary’s prior data and/or opinion), and \( Z \) is the credibility factor, satisfying the condition \( 0 \leq Z \leq 1 \). Under these approaches, the credibility estimator of the quantity of interest is derived as a linear compromise between the current observations and the actuary’s prior opinion. Graphically we see that the compromise estimator, \( C \), is somewhere on the line segment between \( R \) and \( H \), as shown in Figure 1.2.
The symbol $Z$ denotes the weight assigned to the (current) data and $(1-Z)$ the weight assigned to the prior data. This formulation of Equation (1.1), which includes the concept of prior data, is in the spirit of the Bayesian paradigm. As an insurance example, a new insurance rate, $C$, is derived as a weighted average of an old insurance rate, $H$, and an insurance rate, $R$, whose calculation is based solely on observations from a recent period. An alternative interpretation of Equation (1.1) is to let $C$ be the insurance rate for a particular class of business, to let $R$ be the insurance rate whose calculation is based solely on the recent experience of that class, and to let $H$ be the insurance rate whose computation takes into account the experience of all classes combined.

To illustrate the types of practical problems that are addressed by credibility theory, we present here the statement of a problem typical of those solved in this text. Because we have not yet developed the technical machinery required to solve such a problem, we defer its solution until Section 6.6.1 (see Example 6.5).

**Example 1.1**

An insurance company has two policies of group workers’ compensation. The aggregate claim amounts in millions of dollars for the first three policy years are summarized in the table below. Estimate the aggregate claim amount during the fourth policy year for each of the two group policies.

<table>
<thead>
<tr>
<th>Aggregate Claim Amounts</th>
<th>Group Policy</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy Year</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>13</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
Over the years there have been three major approaches to credibility: limited fluctuation, greatest accuracy, and Bayesian. The first two approaches fall under the frequentist paradigm, as neither entails the use of Bayes’ theorem. Moreover, neither approach explicitly requires prior information (i.e., a formal prior probability distribution) in order to compute either the credibility factor, $Z$, or the estimate, $C$. The most well-developed approach to greatest accuracy credibility is least squares credibility. Because this approach was popularized by Hans Buhlmann, it is referred to in this text as Buhlmann’s approach.

1.4 THREE APPROACHES TO CREDIBILITY

The limited fluctuation and Buhlmann approaches both involve the explicit calculation of the credibility factor, $Z$, and the use of Equation (1.1) to obtain the compromise estimator, $C$. On the other hand, the Bayesian approach requires neither the direct calculation of $Z$ nor the use of Equation (1.1).

1.4.1 LIMITED FLUCTUATION APPROACH

Mowbray [1914] described a limited fluctuation approach for deriving the number of exposures required for full credibility, the case where $Z = 1$. Perryman [1932] proposed a limited fluctuation approach to partial credibility problems, those for which $Z < 1$. More modern treatments of the limited fluctuation approach to both full credibility and partial credibility are found in Longley-Cook [1962] and in Chapter 8 of Hossack, Pollard, and Zehnwirth [1983]. Outside of North America this approach is sometimes called “American credibility.”

1.4.2 BÜHLMANN’S APPROACH

Bühlmann’s approach, as described in this text, is based on Bühlmann [1967], which had its origins in a paper by Bailey [1942 and 1943]. Bühlmann and Straub [1972] describe an important generalization of the 1967 Bühlmann work.
1.5 BAYESIAN APPROACH TO CREDIBILITY

1.5.1 BAYESIAN STATISTICAL INFERENCE –
THE EARLY YEARS

The Bayesian approach goes all the way back to the Reverend Thomas Bayes who was born in London, England around 1702. According to Stigler [1986], “Bayes was an ordained Nonconformist minister in Turnbridge Wells (about 35 miles southeast of London).” Although Bayes was elected a fellow of the Royal Society in 1742, his major work was not published until 1764, almost three years after his death. For a long time, his membership in the Royal Society was something of a mystery. Recently-discovered letters, however, now indicate that he did indeed have private correspondence with the other leading intellectuals of his era in London. When Bayes died in 1761, he left £100 and his scientific papers to his friend, Richard Price. After adding an introduction and an appendix, Price presented Bayes’ essay “Toward Solving a Problem in the Doctrine of Chance” to the Royal Society.

The famous French astronomer, probabilist and mathematician Pierre Simon Laplace, who lived from 1749-1827, both championed and extended Bayes’ work. In his text entitled *Essai philosophie sur les probabilités* (*Philosophical Essay on Probabilities*), Laplace described a mathematical framework for conducting statistical inference. This extended the work of Bayes and constituted the essence of Bayesian statistical inference. Laplace took this work seriously as the following passage from the beginning of his Essay indicates:

“Here I will present ... the principles and general results of the theory, applying them to the most important questions of life, which are indeed, for the most part, only questions of probability.”

Inverse Probabilities and Statistical Inference

Bayes’ theorem has practical application in many fields. Kanellos [2003] presents one in a recent article about the application of Bayes’ theorem to data searches entitled “18th Century Theory is New Force in Computing.” Bayes’ Theorem is important to actuaries because it enables them to perform statistical inference by computing inverse probabilities.
What exactly do we mean by “inverse probabilities”? We use the term “inverse” because we are inferring backwards from results (or effects) to causes. Let’s look at some simple examples to examine this further.

A typical probability problem might be stated as follows: I have a standard die with six sides numbered from “one” through “six” and throw the die three times. What is the probability that the result of each of these three tosses of the die will be a “six”?

Now, I might have a second (non-standard) die with three sides numbered “1” and three sides numbered “six.” Again I can ask the same question: What is the probability that the result of each of these three tosses of the die will be a “six”?

The idea behind inverse probabilities is to turn the question around. Here, we might observe that the results of three throws of a die were all “sixes.” We then ask the question: What is the probability that we threw the standard die (as opposed to the non-standard die), given these results?

1.5.2 Whitney’s View of Credibility

Whitney [1918] stated that the credibility factor, $Z$, needed to be of the form

$$Z = \frac{n}{n + k}$$

where $n$ represents “earned premiums” and $k$ is a constant to be determined. The problem was how to determine $k$. Whitney noted that, “In practice $k$ must be determined by judgment.” Whitney also noted that, “The detailed solution to this problem depends upon the use of inverse probabilities” via Bayes’ Theorem.

Predictive Distributions

In insurance work, we typically experience a number of claims or an aggregate amount of losses in one or more prior observation periods. The questions we want to answer are:

---

1 See Whitney [1918, page 289].
2 See Whitney [1918, page 277].
(1) Given such results, how many claims will we experience during the next observation period?

(2) Given such results, what will be the aggregate loss amount during the next observation period?

Using Bayes’ Theorem, we can construct an entire probability distribution for such future claim frequencies or loss amounts. Probability distributions of this type are usually called **predictive distributions**. Predictive distributions give the actuary much more information than would an average or other summary statistic. A predictive distribution provides the actuary with much more information than just the expected aggregate amount of losses in the next period. It provides the actuary with a complete profile of the tail of the probability distribution of aggregate losses for use in a “value-at-risk” analysis. Thus, predictive distributions can provide the actuary and her client an important tool with which to make business decisions under uncertainty.

**1.5.3 Bayesian Statistical Inference and Modern Times**

Perhaps, in part, because the frequentist paradigm of statistics dominated the statistical community during the first half of the twentieth century, it remained for Bailey [1950] to rediscover and advance Whitney’s ideas. During the second half of the twentieth century, Bayesian methods gained increased adherents. Two of the earliest influential books on Bayesian statistics were Savage [1954] and Raiffa and Schlaifer [1961]. Mayerson [1965] brought together the statistical developments in Bayesian statistical inference and the actuary’s credibility problem, reexamining Bailey’s results using the concept of a “conjugate prior distribution” and other more modern notation and terminology. Ericson [1970] and Jewell [1974] generalized Mayerson’s results. Whereas Whitney and Bailey had considered only the distribution of the number of claims, Mayerson, Jones, and Bowers [1968] and Hewitt [1971] considered both the distribution of the number of claims and the distribution of the amount of those claims. Hewitt used some clever, artificial examples to illustrate the use of a full Bayesian approach to insurance ratemaking. It remained for Klugman [1987 and 1992], who had the advantage of modern computing equipment, to extend Hewitt’s ideas and actually apply them to a major practical insurance-ratemaking problem.
1.5.4 BAYESIAN STATISTICAL INFERENCE AND MODERN COMPUTING

With the increased power of 21st-century computing equipment, advances in statistical algorithms (e.g., the EM algorithm and Markov chain Monte Carlo methods) that implement the Bayesian approach, and widely-available software that performs Bayesian inference (i.e., WinBUGS\textsuperscript{3}), a wider class of problems is becoming susceptible to solution via the Bayesian approach.

1.6 EXPLORATORY DATA ANALYSIS

Some of the followers of John Tukey [1977] consider “exploratory data analysis” to be another distinct approach to data analysis\textsuperscript{4}. While it is not the intention here to enter this philosophical discussion, it is often important to do substantial exploratory data analysis prior to constructing formal models, doing statistical inference, or carrying out other types of more involved statistical procedures. There are several reasons for doing this. First, substantial insight can often be gained by using simple approaches. In some situations, especially when the actuary thoroughly understands the subject matter, exploratory data analysis may yield a complete solution. As an example, we consider the following table that summarizes the experience of some mortgages insured by the Federal Housing Administration (FHA) – a component of the U. S. Department of Housing and Urban Development.

\textsuperscript{3} The BUGS (Bayesian inference Using Gibbs Sampling) Project (begun by the MRC Biostatistics unit at Imperial College, London) is concerned with the development of flexible software for Bayesian analysis of complex statistical models using Markov chain Monte Carlo methods. The “Win” prefix refers to Microsoft’s Windows operating system. For more details about BUGS, actuaries should read David Scollnik [2001]: “Actuarial Modeling with MCMC and BUGS.”

\textsuperscript{4} In addition to Tukey’s seminal reference work, cited above, other (perhaps more refined) references on exploratory data analysis include Mosteller and Tukey [1977] and Velleman and Hoaglin [1981].
EXAMPLE 1.2

<table>
<thead>
<tr>
<th>Loan-to-value ratio</th>
<th>Mortgage Amount (In Dollars)</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 80.0%</td>
<td>&lt; 25,000</td>
<td>8.38%</td>
</tr>
<tr>
<td></td>
<td>25,001-35,000</td>
<td>6.88%</td>
</tr>
<tr>
<td></td>
<td>≥ 35,001</td>
<td>6.74%</td>
</tr>
<tr>
<td></td>
<td>≥ 50,001</td>
<td>10.01%</td>
</tr>
<tr>
<td></td>
<td>≥ 60,000</td>
<td>6.94%</td>
</tr>
<tr>
<td></td>
<td>≥ 60,000</td>
<td>7.63%</td>
</tr>
<tr>
<td>80.1 – 85.0</td>
<td>20.43</td>
<td>12.47</td>
</tr>
<tr>
<td>85.1 - 90.0</td>
<td>24.33</td>
<td>17.43</td>
</tr>
<tr>
<td>90.1 – 95.0</td>
<td>27.70</td>
<td>23.53</td>
</tr>
<tr>
<td>95.1 – 97.0</td>
<td>33.48</td>
<td>32.42</td>
</tr>
<tr>
<td>97.1 – 100.0</td>
<td>42.86</td>
<td>52.05</td>
</tr>
</tbody>
</table>

Because we know from a companion table that there are only a small number of mortgages whose loan-to-value ratio is in the 97.1 – 100.0% category, we ignore that line of the table. We find a strong pattern indicating that the claim rate goes down as (1) the mortgage amount goes up and (2) as the loan-to-value ratio goes down. In particular, we note that in the roughly eight years covered by the table, more than 27% of the loans having a loan-to-value ratio in excess of 95% resulted in an insurance claim. The message of this table is clear. If you originate mortgages with little or no down-payment, the proportion of mortgages ending up in foreclosure may be substantial. It does not come as a surprise then that, after lenders originated a large number of mortgages with little or no down-payment during the period 2003-2007, a substantial number of these mortgages ended up in foreclosure. Should it come as a surprise that the housing “bubble” burst?

Second, exploratory data analysis often gives useful insight into the process generating the data. Such insight could be critical to the selection of a good model.

---

5 Claim rate is defined as the proportion of claims received for a given origination year, on single-family mortgages insured by FHA, i.e.,

\[
\text{claim rate} = \frac{\text{number of claims}}{\text{number of mortgages originated}}.
\]
Too often large databases/data warehouses have material deficiencies involving erroneous or missing data elements, missing records, and/or duplicate records. Health insurance companies are concerned with avoiding duplicate claim payments to policyholders. Life insurance companies are concerned with (1) making payments to deceased annuitants and (2) failing to pay beneficiaries of life insurance policyholders because they are not aware that the policyholder has died. Hansen and Wang [1991] describe major deficiencies in a wide range of databases. Thus, the existence of material errors is not an unusual occurrence. Exploratory data analyses can often reveal such errors in the database under study. For a more complete discussion of how to prevent, identify, and correct faulty data, the interested reader should see Herzog, Scheuren, and Winkler [2007].

1.7 RECENT APPLICATIONS

We conclude this chapter by citing some recent applications of credibility theory to actuarial problems. Jewell [1989 and 1990] shows how to use Bayesian procedures to calculate incurred but not yet reported reserve requirements. Russo [1995] extends the work of Jewell. In order to estimate insurance reserves, Russo develops continuous time models of claim reporting and payment processes. In so doing, he employs both the Bayesian paradigm and a multistate model of the incurred claims process.

Klugman [1987] uses a full Bayesian approach to analyze actual data on worker’s compensation insurance. Klugman investigates two problems. First, he calculates the joint posterior distribution of the relative frequency of claims in each of 133 rating groups. He employs three distinct prior distributions and shows that the results are virtually identical in all three instances. Second, Klugman analyzes the loss ratio for three years of experience in 319 rating classes in Michigan. He uses these data to construct prediction intervals for future observations (i.e., the fourth year). He then compares his predictions to the actual results.

The Bayesian paradigm has been used to graduate (or smooth) various types of mortality data. London [1985], building on the pioneering work of Kimeldorf and Jones [1967], provides a general description of this method. London also provides a Bayesian rationale for the historically popular Whittaker graduation method. A specific application of Bayesian graduation is found in Herzog [1983].
Young [1997, 1998] has done some research on credibility and spline functions. Her work enables the actuary to estimate future claims as a function of a statistic other than the sample mean. For example, Young [1998] argues that the use of a regression model with the predictor variable being a function of the sample geometric mean may lead to a more accurate estimator, i.e., one whose squared error loss is reduced.

As discussed in Chapter 8 of this text, Ericson [1970] and Jewell [1974] have shown that the Bühlmann estimate is equal to the Bayesian estimate of the pure premium when the claim distribution belongs to the exponential family of probability distributions and the conjugate prior is employed. Landsman and Makov [1998a] have extended this result to claim distributions belonging to the “exponential dispersion family” of distributions. Landsman and Makov [1998b] suggest a totally new approach to deal with the situation in which the claim distribution is not a member of either of the two previously-mentioned families of distributions.

Frees, et al., [1999] and Frees, et al., [2001] delineate the relationship between (1) credibility models and (2) parametric statistical models used for panel (longitudinal) data analysis.

Prior to the advent of Markov Chain Monte Carlo (MCMC) numerical methods, it was only feasible to implement a full Bayesian approach for a limited class of models. Scollnik [2001] shows how to implement Bayesian methods in actuarial models using the BUGS software package. Fellingham, Tolley, and Herzog [2005] also use BUGS to construct a Bayesian hierarchical model in order to estimate health insurance claim costs. Finally, Rosenberg and Farrell [2008] use version 1.4 of WinBUGS to construct a Bayesian statistical model in order to predict the incident and cost of hospitalization for a group of children with cystic fibrosis.
1.8 Exercises

1.1 Introduction

1-1 According to Rodermund, what has been the casualty actuaries’ most important and enduring contribution to casualty actuarial science?

1.2 Statistical Paradigms

1-2 Name the two major statistical paradigms of current interest.

1.3 What Is Credibility?

1-3 Using Equation (1.1), determine the realization of the compromise estimator $C$, given that (i) the mean of the current observations is 10, (ii) the prior mean is 6, and (iii) the credibility factor is .25.

1-4 Using Equation (1.1), determine the insurance rate, $C$, for a particular class of business given that (i) the insurance rate calculated strictly from the experience data of that class of business is $100, (ii) the insurance rate for all classes combined is $200, and (iii) the credibility factor for the class is .40.

1.4 Three Approaches to Credibility

1-5 List the three major approaches to credibility.