PART I

Foundations and Principles

CHAPTER 1

Background

1.1 RATIONALE FOR BAYESIAN INFERENCE AND PRELIMINARY VIEWS OF BAYES' THEOREM

In 1763 an important scientific paper was published in England, authored by a Reformist Presbyterian minister by the name of Thomas Bayes (Bayes, 1763). The implications of the paper indicated how to make statistical inferences that build upon earlier understanding of a phenomenon, and how *formally* to combine that earlier understanding with currently measured data in a way that updates the degree of belief (subjective probability) of the experimenter. The earlier understanding and experience is called the "prior belief" (belief or understanding held prior to observing the current set of data, available either from an experiment or from other sources), and the new belief that results from updating the prior belief is called the "posterior belief" (the belief held after having observed the current data, and having examined those data in light of how well they conform with preconceived notions). This inferential updating process is eponymously called Bayesian inference. The inferential process suggested by Bayes shows us that to find our subjective probability for some event, proposition, or unknown quantity, we need to multiply our prior beliefs about the event by an appropriate summary of the observational data. Thus, Bayesian inference suggests that all formal scientific inference inherently involves two parts, a part that depends upon subjective belief and scientific understanding that the scientist has prior to carrying out an experiment, and a part that depends upon observational data the scientist obtains from the experiment. We present Bayes' theorem compactly here in order to provide an early insight into the development of the book in later chapters.

Briefly, in its most simple form, the form for events (categorical or discrete data), Bayes' theorem or formula asserts that if $P\{A\}$ denotes the probability of an event A, and $P\{B|A\}$ denotes the probability of an event B conditional on knowing A, then:

$$P\{B|A\} = \frac{P\{A|B\}P\{B\}}{P\{A|B\}P\{B\} + P\{A|\bar{B}\}P\{\bar{B}\}},$$

where $\bar{B}$ denotes the complementary event to event B. This simple statement of conditional probability is the basis for all Bayesian analysis. $P(B)$ denotes the prior belief about $B, P\{B|A\}$ denotes the posterior belief about B (once we know A), and $P\{A|B\}$ denotes the model, that is, the process that generates the event A based upon knowing B.

As an example, suppose you take a laboratory test for diabetes. Let A denote the outcome of the test; it is a *positive* outcome if the test finds that you have the tell-tale indicators of diabetes, and it is a *negative* outcome if you do not. But do you really have the disease? Sometimes, although you do not actually have diabetes, the test result is positive because of imperfect characteristics of the laboratory test. Similarly, sometimes when you take the test there is a negative outcome when in fact you do have the disease. Such results are called false positives and false negatives, respectively. Let B denote the event that you actually do have diabetes. You would like to know the chances that you have diabetes in light of your positive test outcome, $P\{B|A\}$. You can check with the laboratory to determine the sensitivity of the test. Suppose you find that when the test is negative, the error rate is 1 percent (false negative error rate), and when the test is positive, its accuracy is 3 percent (the false positive error rate). In terms of the Bayes' formula,

$$\begin{aligned} P\{A = +test|B = diabetes\} &= 1 - P\{\bar{A} = -test|B = diabetes\} \\ &= 1 - 0.01 = 0.099, \end{aligned}$$

and

$$P\{+test|\bar{B} = no\, diabetes\} = \text{probability of a false positive} = 0.03.$$

Bayes' formula then gives:

$$P\{B\,|+\} = \frac{P\{+|B\}P\{B\}}{P\{+|B\}P\{B\} + P\{+\,|\bar{B}\}P\{\bar{B}\}}$$

$$P\{diabetes\,|+\} = \frac{(0.99)P\{B\}}{(0.99)P\{B\} + (0.03)P\{B\}}.$$

It only remains to determine $P\{B\}$, the chances of someone having diabetes. Suppose there is no indication that diabetes runs in your family, so the chance of you having diabetes is that of a randomly selected person in the population about your age, say about one chance in one million, that is, $P\{B\} = 10^{-6}$. Substituting in the above formula gives:

$$P\{\,you\ have\ diabetes\,|positive\ test\ result\} = 0.003 = 0.3\%.$$

If we are concerned with making inferences about an unknown quantity θ, which is continuous, Bayes' theorem takes the form appropriate for continuous θ:

$$h(\theta|x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n \mid \theta)g(\theta)}{\int f(x_1, \ldots, x_n \mid \theta)g(\theta)d\theta},$$

where $h(\cdot)$ denotes the probability density of the unknown θ subsequent to observing data $(x_1, \ldots, x_n)$ that bear on θ, f denotes the likelihood function of the data, and g denotes the probability density of θ prior to observing any data. The integration is taken over the support of θ. This form of the theorem is still just a statement of conditional probability, as we will see in Chapter 4.

A large international school of scientists (some of whom even preceded Bayes) supported, expanded, and developed Bayesian thinking about science. These include such famous scientists as James Bernoulli, writing in 1713, Pierre Simon de Laplace, writing in 1774, and many nineteenth- and twentieth-century scientists. Today, scientists schooled in the Bayesian approach to scientific inference have been changing the way statistical methodology itself has been developing. Many believe that a paradigm shift has been taking place in the way scientific inference is carried out, away from what is sometimes referred to as classical, or frequentist, statistical inference. Many scientists now recognize the advantages of bringing prior beliefs into the inferential process in a formal way from the start, instead of striving, and almost inevitably failing, to achieve total objectivity, and bringing the prior information into the problem anyway, in surreptitious, or even unconscious ways. Subjectivity may enter the scientific process surreptitiously in the form of seemingly arbitrarily imposed constraints in the introduction of initial and boundary conditions in the arbitrary levels of what should be called a significant result (selecting the "level of significance"), and in the de-emphasizing of certain outlying data points that represent suspicious observations.

Scientists will see that Bayes' theorem gives the degree of a person's belief (that person's subjective probability) about some unknown entity once something about it has been observed (i.e., posterior to collecting data about that entity), and shows that this subjective probability is proportional to the product of two types of information. The first type of information characterizes the data that are observed; this is usually thought of as the objective portion of posterior belief, since it involves the collection of data, and data are generally thought to be objectively determined. (We recognize that we do not really mean that data are objective unless we assume that there were no subjective influences surrounding the data collected.) This so-called *objective information* is summarized in the likelihood function. But the likelihood function is of course almost invariably based upon data that has been influenced by the subjectivity of the observer. Moreover, in small or often in even moderate size samples its structural form is not very well determined. So the likelihood function will almost invariably contain substantial subjective influences and uncertainty.

The second type of information used in Bayesian analysis is the person's degree of belief, the subjective probability about the unknown entity, held prior to observing anything related to it. This belief may be based, at least in part, on things that were observed or learned about this unknown quantity prior to this most recent measurement. Using Bayes' theorem, scientific belief about some phenomenon is formally updated by new measurements, the idea being that we learn about something by modifying what we already believe about it (our prior belief) to obtain a posterior belief after new observations are taken.

While it is well known that for a wide variety of reasons there are always some subjective influences in the research of scientists, and always have been, it is less well known that strong major subjective influences have actually been present in some of the work of the most famous scientists in history (see, for example, Press and Tanur, 2001). The personal beliefs and opinions of these scientists have often very strongly influenced the data they collected and the conclusions they drew from those data. While the phenomena these scientists were investigating were generally truly objective phenomena, external to the human mind, nevertheless, the data collected about these phenomena, and the decisions made relative to these phenomena were often driven by substantial subjectivity. Bayesian analysis, had it been available to these scientists, and had it been used, might have permitted these scientists to distinguish between models whose coexistence has caused controversy about their results even hundreds of years later.

Further, several scientists examining the same set of data from an experiment often develop different interpretations. This phenomenon is not unusual in science. When several scientists interpret the same set of data they rarely have *exactly* the same interpretations. Almost invariably, their own prior beliefs about the underlying phenomenon enter their thinking, as do their individual understanding of how meaningful each data point is. Their conclusions regarding the extent to which the data support the hypothesis will generally reflect a mixture of their prior degree of belief about the hypothesis they are studying, and the observed data.

Thus, we see that whether formal Bayesian inference is actually used in dealing with the data in an experiment, or whether other, nonBayesian methods are used, subjective prior belief is used in one way or another by all good scientists in a natural, and sometimes quite informal, way. Science cannot, and should not, be totally objective, but should and does involve a mixture of both subjective and objective procedures, with the one type of procedure feeding back on the other. As the data show the need for modification of the hypothesis, a new hypothesis is entertained, a new experiment is designed, new data are taken, and what was posterior belief in the earlier experiment becomes the prior belief in the new experiment, because the result of the last experiment is now the best understanding the scientist has of what result to expect in a new experiment. To study the future, scientists must learn from the past, and it is important—indeed inevitable—that the learning process be partly subjective.

During the twentieth century, since the development of methods of Bayesian statistical inference, there have been many exciting new scientific discoveries and developments. Some have been simply of the qualitative type where certain phenomena have been discovered that were not previously known (such as the discovery of the existence of the radiation belts that surround the Earth, the discovery of superconductivity, or the discovery of the double helical structure of DNA), and others have been quantitative, establishing relationships not previously established (such as the discoveries of the dose/effect relationships of certain pharmaceutical drugs, vaccines, and antibiotics that would minimize the chances of contracting various infectious diseases, or maximize the chance of a cure).

Considerable scientific advance is based upon finding important phenomena that are sometimes so shrouded in noise that it is extremely difficult to distinguish the phenomenon of interest from other factors and variables. In such cases, prior information about the process, often based upon previous theory, but sometimes on intuition or even wild guesses, can often be profitably brought to bear to improve the chances of detecting the phenomenon in question. A considerable amount of Bayesian statistical inference procedures that formally admit such prior information in the scientific process of data analysis have had to await the advent of modern computer methods of analysis, an advent that did not really occur until the last couple of decades of the twentieth century. However, since the arrival of real-time interactive computers, computational Bayesian methods such as Markov Chain Monte Carlo (MCMC, see Chapter 6) have been very usefully applied to problems in imaging and other problems in physics and engineering (see the series of books edited by different authors, every year since 1980, Maximum Entropy and Bayesian Methods published by Kluwer), problems of meta-analysis to synthesize results in a field—in biology, medicine, economics, physics, sociology, education, and others—and in a variety of scientific fields (see, for example, Appendix 5).

Subjectivity in science implies that we generally arrive at universal scientific truths by a combination of subjective and objective means. In other words, the methodology we use to discover scientific truths benefits greatly from bringing informed scientific judgment to bear on the hypotheses we formulate, and on the inferences we make from data we collect from experiments designed to test these hypotheses. Informed scientific judgment should not be shunned as a nonobjective, and therefore a poor methodological approach; collateral information about the underlying process should be actively sought so that it can be used to improve understanding of the process being studied. Combining informed knowledge with experimental data will generally improve the accuracy of predictions made about future observations.

Subjectivity is an inherent and required part of statistical inference and the scientific method. It is a *sine qua non* in the process of creating new understanding of nature. It must play a fundamental role in how science is carried out.

However, excessive, informal, untested subjectivity in science is also responsible for some basic errors, misrepresentations, overrepresentations, or scientific beliefs that were later shown to be false, that have occurred in science (see, for example, Press and Tanur, 2001). This author's views of subjectivity in science coincide closely with those of Wolpert (1992, p. 18) who wrote:

> . . . the idea of scientific objectivity has only limited value, for the way in which scientific ideas are generated can be highly subjective, and scientists will defend their views vigorously . . . It is, however, an illusion to think that scientists are unemotional in their attachment to their scientific views—they may fail to give them up even in the face

> of evidence against them . . . scientific theories involve a continual interplay with other scientists and previously acquired knowledge . . . and an explanation which the other scientists would accept.

To illustrate the notion that subjectivity underlies experimental science, in Section 1.2 we use a very simple example involving whether or not a desired effect is observed in an experiment to show that merely observing scientific data and forming a likelihood function can involve considerable subjectivity.

1.2 EXAMPLE: OBSERVING A DESIRED EXPERIMENTAL EFFECT

Let us suppose that 100 observations are collected from an experiment replicated 100 times; there is one observation from each replication. These data are sent to five scientists located in five different parts of the world. All five scientists examine the same data set, that is, the same 100 data points. (Note that for the purposes of this example, the subjectivity involved in deciding what data to collect and in making the observations themselves is eliminated by sending the same "objective" data to all five scientists.) Should we expect all five of the scientists to draw the same conclusions from these data?

The answer to this question is a very definite "no". But how can it be that different observers will probably draw different conclusions from precisely the same data? As has been said above, inferences from the data will be a mixture of both subjective judgment (theorizing) and objective observation (empirical verification). Thus, even though the scientists are all looking at the same observational data, they will come to those same data with differing beliefs about what to expect. Consequently, some scientists will tend to weight certain data points more heavily than others, while different scientists are likely to weight experimental errors of measurement differently from one another. Moreover, if scientists decide to carry out formal checks and statistical tests about whether the phenomenon of interest in the experiment was actually demonstrated (to ask how strongly the claimed experimental result was supported by the data), such tests are likely to have different results for different scientists, because different scientists will bring different assumptions to the choice of statistical test. More broadly, scientists often differ on the mathematical and statistical models they choose to analyse a particular data set, and different models usually generate different conclusions. Different assumptions about these models will very often yield different implications for the same data.

These ideas that scientists can differ about the facts are perhaps startling. Let us return to our 100 observations and five scientists to give a very simple and elementary example, with the assurance that analogous arguments will hold generally for more realistic and more complicated situations.

Let us assume that the purpose of the experiment is to determine the probability that a certain genetic effect will take place in the next generation of a given type of simple organism. The question at issue is whether the effect occurs randomly or is subject to certain genetic laws. If the experiment is carried out many times, independently, under approximately the same set of conditions, how frequently will the genetic effect be observed? If the effect is observed say, 50 percent of the time, we will say it occurs merely at random, but it might occur, say 60 percent of the time, or more frequently. More generally, what percentage of the time will it occur under those experimental conditions? Call the (unknown) proportion of the time the effect should be observed, when the experiment is carried out repeatedly and independently, p; $p = 0.5$ if the effect is merely observed at random.

To determine p you repeat the experiment many times, each time record the result, and on the basis of the outcome, determine whether the effect is observed (called a "success"), or not (called a "failure"). In our case the experiment is repeated 100 times, and the outcome of success or failure is recorded for each trial of the experiment. Next, suppose that 90 successes were obtained. The question now is, "What is the value of p"?

These are the data you send to the five scientists in five different locations around the world (three women and two men) to see how they interpret the results. You tell each of them that there were 90 successes out of 100 trials of an experiment (or 90 percent of the experiments had successful outcomes). You also tell them that you plan to publish their estimates and their reasoning behind their estimates in a professional scientific journal, and that their reputations are likely to be enhanced or to suffer, in proportion to their degrees of error in estimating p. As we shall see, it will turn out that they will all have different views about the value of p after having been given the results of the experiment.

Scientist #1 is largely a *theorist* by reputation. She thinks successes are occurring at random since her theoretical understanding of the biochemical mechanism involved suggests to her that there should not be an effect; so for her, $p = 0.5$ no matter what. Her line of reasoning regarding the experimental outcomes is that although it just happened that 90 percent of the first 100 replications of the experiment were successful, it does not mean that if the experiment were to be repeated for another 100 trials, the next 100 trials would not produce, say, 95 failures, or any other large proportion of failures. Scientist #1 has a very strong preconceived belief based upon theory that the effect should not take place ($p = 0.5$), in the face of real data that militates against that belief. For her, unless told otherwise, all such experiments under these preset conditions should not demonstrate any real effect, even if many runs of successes or many runs of failures just happen to occur.

Scientist #2 has the reputation for being an *experimentalist*. He thinks $p = 0.9$, because that is the proportion of successes found in 100 replications. (This estimate of p is actually the maximum likelihood estimate.) Scientist #2's definition of the best estimate available from the data is the fraction of successes actually obtained. While Scientist #1 believed strongly in theory, Scientist #2 is ready to abandon theory in favor of strong belief in data, regardless of theory.

Scientist #3 is also a well-known *skeptic*. She decides that there is something strange about the reported results since they violate her strongly held expectation that the effect really should not be observed, other than at random, so there should be only about 50 successes in 100 replications. Scientist #3 then writes to you and asks you for details about the experimental equipment used. When Scientist #3 receives these details, she decides to replicate the experiment herself another 100 times. Most of the time she finds successes. These efforts convince Scientist #3 that the effect is really being produced, just as Scientist #2 concluded. The effect actually turns up 82 times. But what should Scientist #3 do now? Since her recent collection of replications found 82 successes in her new 100 trials, and the previous collection found 90 successes, Scientist #3 reasons that the experiment has been replicated 200 times, so perhaps it would not be unreasonable to estimate p as $(82 + 90)/200 = 0.86$.

When he actually carries out experiments in his research, which is only occasionally, Scientist #4 is an extremely *thorough and careful experimentalist*. He feels that he should know more about the experiment outcomes than he has been told so far. He writes to you for a copy of the actual sequence of experiment outcomes, studies it, and decides to ignore a run of 50 straight successes that occurred, feeling that such a run must be a mistake in reporting since it is so unlikely. This reduces the sample size (the number of available data points) from 100 down to 50, and out of those 50, 40 were successes. So his conclusion is that $p = 40/50 = 0.8$. Scientist #4 has taken the practical posture that many scientists take of weighting the observations so that some observations that are believed to be errors are discarded or downweighted in favor of those thought to be better measurements or more valid in some experimental sense.

Scientist #5 may be said to be *other-directed*. She badly wants the recognition from her peers that you said would depend on the accuracy of her estimate. She learns from you the names of the other four scientists. She then writes to them to find out the estimates they came up with. Having obtained their results, Scientist #5 decides that the best thing to do would be to use their average value. So for Scientist #5, the estimate of $p = (0.5 + 0.9 + 0.86 + 0.8)/4 = 0.765$. Or perhaps, learning that Scientist #1 made an estimate of p that was not data dependent, Scientist #5 might eliminate Scientist #1's estimate from her average, making the subjective judgment that it was not a valid estimate. Then Scientist #5 would have an estimate of $p = (0.9 + 0.86 + 0.8)/3 = 0.853$. Scientist #5's strategy is used by many scientists all the time so that the values they propose will not be too discrepant with the orthodox views of their peers.

So the five scientists came up with five different estimates of p for the same observed data. All the estimates are valid. Each scientist came to grips with the data with a different perspective and a different belief about p.

Suppose there had been a sixth scientist, a *decision theorist*, driven by a need to estimate unknown quantities on the basis of using them to make good decisions. Such a scientist would be interested in minimizing the costs of making mistakes, and might perhaps decide that overestimating p is as bad as underestimating it, so the costs should be the same for making these two types of errors. Moreover, he wants to select his estimator of p to enhance his reputation, and you have told him that a correct estimate will do just that. So he decides to select his estimator in such a way that the cost of being wrong will be as small as possible—regardless of the true value of p (under such circumstances he would often adopt a "quadratic loss function"). If Scientist #6 were to adopt the subjective belief that all values of p are equally likely, and then he used Bayes' theorem, his resulting estimate of p would be $p = 91/102 = 0.892$.

In summary, the values of p found by the six scientists are thus:

Scientist	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**
Estimated value of *p*	0.500	0.900	0.860	0.800	0.765 or 0.853	0.892

But what is the true value of p? Note that with the exception of Scientist #1, who refuses to be influenced by the data, all the scientists agree that p must be somewhere between 0.765 and 0.900, agreeing that the effect is not occurring at random, but disagreeing about how often it should occur theoretically.

Suppose the experiment had been replicated 1000 times instead of the 100 times we have just considered, but with analogous results of 90 percent successes. Would this have made any difference? Well perhaps it might have to Scientist #1, because people differ as to the point at which they will decide to switch positions from assuming the effect is occurring only at random, in spite of the experimental outcome results, to a position in which they are willing to assume the effect is actually being generated from the experimental conditions. One scientist may switch after 9 successes out of 10 trials, another after 90 out of 100, whereas another may not switch until there are perhaps 900 successes out of 1000 trials, or might insist on an even more extensive experiment. In any case the scientists may still differ in their views about the true value of p for a very wide variety of reasons, only a few of which are mentioned above.

The biochemical experiment example just discussed was very elementary as far as experiments go, but it was science nevertheless. Similar interpretive and methodological issues arise in all branches of the sciences. In this book we examine the probabilistic and inferential statistical foundations, principles, methods, and applications of Bayesian statistical science, and throughout we adopt the notion of subjective probability, or degree of belief.

1.3 THOMAS BAYES

Thomas Bayes was a Presbyterian minister and mathematician who lived in England in the 1700s (born about 1702 and died April 17, 1761). Richard Price, a friend of Bayes, and interested in Bayes' research, submitted Bayes' manuscript on inverse probability in the binomial distribution to the professional journal, *Philosophical Transactions of the Royal Society*, which published the paper (posthumously) in 1763, an article reproduced in Appendix 4 of this book, along with biographical information about Bayes reproduced in Appendices 1 to 3.

There has been some mystery associated with Thomas Bayes. We are not quite certain about the year of his birth or about the authenticity of his portrait. Moreover, questions have been raised about what his theorem actually says (Stigler, 1982), and who actually wrote the paper generally attributed to him (Stigler, 1983). The common interpretation today is that the paper proposed a method for making probability inferences about the parameter of a binomial distribution conditional on some observations from that distribution. (The theorem attributed to Thomas Bayes is given in Proposition 9, Appendix 4; the Scholium that follows it has been controversial, but has been widely interpreted to mean that "knowing nothing" about a parameter in the unit interval implies we should take a uniform distribution on it.) Common belief is that Bayes assumed that the parameter had a uniform distribution on the unit interval. His proposed method for making inferences about the binomial parameter is now called Bayes' theorem (see Section 2.2) and has been generalized to be applicable beyond the binomial distribution, to any sampling distribution. Bayes appears to have recognized the generality of his result but elected to present it in that restricted binomial form. It was Laplace (1774) who stated the theorem on inverse probability in general form, and who, according to Stigler (1986), probably never saw Bayes' essay, and probably discovered the theorem independently. (Bayes carried out his work in England, where his theorem was largely ignored for over 20 years; Laplace carried out his work in France.) Jeffreys (1939) rediscovered Laplace's work.

Your distribution for the unknown; unobservable parameter is called *your* prior distribution, because it represents the distribution of *your* degree of belief about the parameter prior to *your* observing new data, that is, prior to your carrying out a new experiment that might bear on the value of the parameter. Bayes' theorem gives a mathematical procedure for updating your prior belief about the value of the parameter to produce a posterior distribution for the parameter, one determined subsequent to your having observed the outcome of a new experiment bearing on the value of the unknown parameter. Thus, Bayes' theorem provides a vehicle for changing, or updating, the degree of belief about an unknown quantity (a parameter, or a proposition) in light of more recent information. It is a formal procedure for merging knowledge obtained from experience, or theoretical understanding of a random process, with observational data. Thus, it is a "normative theory" for learning from experience, that is, it is a theory about how people *should* behave, not a theory about how they *actually do* behave, which would be an "empirical theory."

The ideas in the theorem attributed to Bayes were really conceived earlier by James Bernoulli in 1713, in Book 4 of his famous treatise on probability, *Ars Conjectandi* ("The Art of Conjecturing"), published posthumously. In that book, James Bernoulli, or Jakob Bernoulli as he was known in German, not only developed the binomial theorem and laid out the rules for permutations and combinations but also posed the problem of inverse probability of Bayes (who wrote his essay 50 years later). However, Bernoulli did not give it mathematical structure. In *Ars Conjectandi*, James Bernoulli (1713) wrote:

> To illustrate this by an example, I suppose that without your knowledge there are concealed in an urn 3000 white pebbles and 2000 black pebbles, and in trying to determine the numbers of these pebbles you take out one pebble after another (each time replacing the pebble you have drawn before choosing the next, in order not to decrease the number of pebbles in the urn), and that you observe how often a white and how often a black pebble is withdrawn. The question is, can you do this so often that it

> becomes ten times, one hundred times, one thousand times, etc., more probable (that is, it be morally certain) that the numbers of whites and blacks chosen are in the same 3 : 2 ratio as the pebbles in the urn, rather than in any other different ratio?

This is the problem of inverse probability, which concerned eighteenth-century mathematicians (Stigler, 1986, Chapter 2). According to Egon Pearson (1978, p. 223), James Bernoulli "...was destined by his father to be a theologian, and [he] devoted himself after taking his M.A. at Basel (Switzerland) to theology." This endeavor involved Bernoulli in philosophical and metaphysical questions (Bayes, of course, received similar training, being a minister). In fact, Maistrov (1974, p. 67), evaluating Bernoulli, believes he was an advocate of "metaphysical determinism," a philosophy very similar to that of Laplace, some of whose work did not appear until 100 years after Bernoulli. Writing in *Ars Conjectandi*, in the first chapter of the fourth part, Bernoulli said:

> For a given composition of the air and given masses, positions, directions, and speed of the winds, vapor, and clouds and also the laws of mechanics which govern all these interactions, tomorrow's weather will be no different from the way it should actually be. So these phenomena follow with no less regularity than the eclipses of heavenly bodies. It is, however, the usual practice to consider an eclipse as a regular event, while (considering) the fall of a die, or tomorrow's weather, as chance events. The reason for this is exclusively that succeeding actions in nature are not sufficiently well known. And even if they were known, our mathematical and physical knowledge is not sufficiently developed, and so, starting from initial causes, we cannot calculate these phenomena, while from the absolute principles of astronomy, eclipses can be precalculated and predicted... The chance depends mainly upon our knowledge.

In the preceding excerpt, Bernoulli examined the state of tomorrow's weather, given today's observational data that relate to weather and a belief about tomorrow's weather. He noted that, of necessity, because of our inability to understand precisely the behavior of the forces governing weather, we must treat tomorrow's weather as uncertain and random to some extent, but predictable in terms of chance (probability), in accordance with our knowledge. This is precisely the kind of question addressable by Bayes' theorem of 1763, in terms of degree of belief about tomorrow's weather, given today's observations and a prior belief about tomorrow's weather. Moreover, the development of quantum mechanics and the Heisenberg uncertainty principle have elaborated Bernoulli's view of chance, showing that chance is a fundamental property of nature that goes beyond mere lack of knowledge of the physical laws governing some physical phenomenon.

Additional background on the life of Thomas Bayes may be found in Appendices 1 to 3.

1.4 BRIEF DESCRIPTIONS OF THE CHAPTERS

This book is subdivided into four parts. Part 1 includes Chapters 1 to 5 on foundations and principles, Part 2 includes Chapters 6 and 7 on numerical implementation of the Bayesian paradigm, Part 3 includes Chapters 8 to 11 on Bayesian inference and decision making, and Part 4 includes Chapters 12 to 16 on models and applications.

In this introductory chapter we have shown that the interpretation of scientific data has always involved a mixture of subjective and objective methods. We have pointed out that it is no longer necessary to adopt the informal subjective methods used by scientists in the past, but that now, Bayesian inference is available to formalize the introduction of subjectively based information into the analysis of scientific data, and computational tools have been developed to make such inference practical. Chapter 2 discusses probability as a degree of belief, calibration, and the axiomatic basis of Bayesian inference and decision making. Chapter 3 presents the meaning of the likelihood principle and conditional inference. Chapter 4, the central chapter in the book, presents and interprets Bayes' theorem for both discrete and continuous data and discrete and continuous parameters. Chapter 5 presents a wide variety of subjective and objective prior distributions, both discrete and continuous, for individuals and for groups, in both univariate and multivariate parameter cases.

Chapters 6 and 7 focus on computer routines that sample from the posterior distribution, approximations, and relevant software. We treat Gibbs sampling, the Metropolis–Hastings algorithm, data augmentation, and some computer software that has been developed for use in Bayesian analysis. In Complement A to Chapter 6 we have provided a detailed discussion of the WinBUGS Program, probably the most widely used software program that affords access to the Gibbs sampling paradigm without doing your own programming. Complement B to Chapter 6 provides a listing of other popular Bayesian software. Chapter 7 discusses the role of large sample posterior distributions (normal), the Tierney–Kadane and Naylor–Smith approximations, and importance sampling.

Chapter 8 discusses Bayesian estimation, both point and interval, and Chapter 9 discusses hypothesis testing. Both treat univariate and multivariate cases. Chapter 10 discusses predictivism, including the de Finetti theorem, de Finetti transforms, and maximum entropy. Chapter 11 presents the Bayesian approach to decision making. In the last part of the volume, in Chapter 12, we discuss the Bayesian approach to the analysis of the general linear model, both univariate and multivariate, including regression, analysis of variance and covariance, and multivariate mixed models. In Chapter 13 we discuss Bayesian model averaging to account for model uncertainty. Chapter 14 explicates Bayesian hierarchical modeling, Chapter 15 discusses Bayesian factor analysis, and Chapter 16 concludes with a presentation of Bayesian classification and discrimination methods.

There are seven appendices, the first three of which relate to the life of Thomas Bayes, and the posthumous submission of his famous theorem to a journal. Appendix 4 presents Bayes' original paper in its entirety. Appendix 5 is a list of references, by subject, of applications of the Bayesian paradigm that have been made across diverse disciplines. It is not intended to be an exhaustive list, merely an indication of a few of the applications that have been made. Appendix 6 presents an explanation of how the most important contributors to the development of Bayesian statistics, a Bayesian Hall of Fame, were selected. There are exercises at the end of each chapter. Appendix 7 provides solutions to selected exercises.

SUMMARY

In this chapter we have examined the role of subjectivity in science and have found that subjectivity, judgment, and degree of belief are fundamental parts of the scientific process. We have also seen from the example involving observing a desired experimental effect (Section 1.2) that observing some phenomenon involves an interpretation of the observational data. Different observers will often have different interpretations of the same data, depending upon their own backgrounds, beliefs they brought to the experiment *a priori*, and theoretical expectations of the experimental outcomes.

EXERCISES

1.1 Give examples of subjectivity in medical diagnosis by a physician; in the introduction of "evidence" by an attorney in civil or criminal trials; in studies of the human genome by molecular biologists; in portfolio analysis of financial securities by certified financial analysts; in the design of a building by a civil engineer; in the "reading" of X rays by a radiologist; and in the evaluation of objects found at an ancient archeological site.

1.2* By going to the original reference sources, explain who Stigler believes actually wrote Bayes' theorem, and also, explain what Stigler believes Bayes' theorem was really saying.

1.3 In the observing of a desired experimental effect example of Section 1.2 of this chapter, can you think of other interpretations of the same data?

1.4* Why is the book *Ars Conjectandi* by James Bernoulli of particular interest to people interested in Bayesian statistical inference?

1.5 In the observing of a desired experimental effect example of Section 1.2 of this chapter, which interpretation of the data seems most appropriate to you? Why?

1.6 In an experiment you were doing, or one that you were analysing, how would you decide whether certain observations were too large or too small to be considered as part of the basic data set?

1.7* Distinguish between prior and posterior distributions.

1.8* The methodological approaches to science of some of the most famous scientists in history (Aristotle, Marie Curie, Charles Darwin, Albert Einstein, Michael Faraday, Sigmund Freud, Galileo Galilei, William Harvey, Antoine Lavoisier, Isaac Newton, Louis Pasteur, and Alexander von Humboldt) were sometimes quite subjective. Give examples in the cases of three of these people. (*Hint*: see Press and Tanur, 2001.)

1.9* Explain the meaning of the statement, "Most scientific inference is already partly subjective, and it always has been."

* Solutions for asterisked exercises may be found in Appendix 7.

FURTHER READING

Bayes, T. (1763). "An Essay Towards Solving a Problem in the Doctrine of Chances," *Philos. Trans. Royal Soc. London*, **V61.53**, 370–418. Reprinted in Appendix 4.

Bernoulli, J. (1713). *Ars Conjectandi*, Book 4, Baseae, Impensis Thurnisiorum.

Gatsonis, C., Hodges, J. S., Kass, R. E., and Singpurwalla, N. D., Eds. (1993). *Case Studies in Bayesian Statistics*, New York, Springer-Verlag.

Gatsonis, C., Hodges, J. S., Kass, R. E., and Singpurwalla, N. D., Eds. (1995). *Case Studies in Bayesian Statistics*, Vol. II, New York, Springer-Verlag.

Gatsonis, C., Hodges, J. S., Kass, R. E., McCulloch, R., Rossi, P., and Singpurwalla, N. D. (1997). *Case Studies in Bayesian Statistics*, Vol. III, New York, Springer-Verlag.

Gatsonis, C., Kass, R. E., Carlin, B., Carriquiry, A., Gelman, A., Verdinelli, I., and West, M. (1999). *Case Studies in Bayesian Statistics*, Vol. IV, New York, Springer-Verlag.

Howson, C. and Urbach, P. (1989). *Scientific Reasoning*, La Salle, Illinois, Open Court Publishing Co.

Jeffreys, H. (1939, 1st edition; 1948, 2nd edition; 1961, 3rd edition). *Theory of Probability*, Oxford, Oxford University Press and Clarendon Press.

Laplace, P. S. (1774, 1814). *Essai Philosophique sur les Probabilites*, Paris. This book went through five editions (the fifth was in 1825) revised by Laplace. The sixth edition appeared in English translation by Dover Publications, New York, in 1951. While this philosophical essay appeared separately in 1814, it also appeared as a preface to his earlier work, *Theorie Analytique des Probabilites*.

Maistrov, L. E. (1974). *Probability Theory: A Historical Sketch*, S. Kotz, Trans. and Ed. New York, Academic Press.

Matthews, R. A. J. (1998a). "Fact versus Factions: the Use and Abuse of Subjectivity in Scientific Research", ESEF Working Paper 2/98, *The European Science and Environment Forum*, Cambridge, England, September, 1998.

Matthews, R. A. J. (1998b) "Flukes and Flaws," *Prospect*., 35, 20–24.

Pearson, E. (1978). *The History of Statistics in the 17th and 18th Centuries*, New York, Macmillan.

Press, S. J. and Tanur, J. M. (2001). *The Subjectivity of Scientists and the Bayesian Approach*, New York, John Wiley and Sons, Inc.

Stigler, S. M. (1982). "Thomas Bayes and Bayesian Inference," *J. Roy. Statist. Soc.* (*A*), **145** (2) 250–258.

Stigler, S. M. (1983). "Who Discovered Bayes' Theorem," *Amer. Stat*., **37**(4), 290–296.

Stigler, S. M. (1986). *The History of Statistics*, Cambridge, MA, The Belknap Press of Harvard University Press.

Wolpert, L. (1992). *The Unnatural Nature of Science*, Boston, Harvard University Press.