Probability and Statistics for Business Decisions

AN INTRODUCTION TO MANAGERIAL ECONOMICS UNDER UNCERTAINTY

Robert Schlaifer
PROFESSOR OF BUSINESS ADMINISTRATION
HARVARD UNIVERSITY

McGRAW-HILL BOOK COMPANY
New York Toronto London
1959

 Library of Congress Catalog Card Number 58-13017

10 11 12 13 14 15 16 - MP - 1 0 9 8

55309

Probability and Statistics for Business Decisions

Preface

This book is a nonmathematical introduction to the logical analysis of practical business problems in which a decision must be reached under uncertainty. The analysis which it recommends is based on the modern theory of utility and what has come to be known as the "personal" definition of probability; the author believes, in other words, that when the consequences of various possible courses of action depend on some unpredictable event, the *practical* way of choosing the "best" act is to assign values to consequences and probabilities to events and then to select the act with the highest expected value. In the author's experience, thoughtful businessmen intuitively apply exactly this kind of analysis in problems which are simple enough to allow of purely intuitive analysis; and he believes that they will readily accept its formalization once the essential logic of this formalization is presented in a way which *can* be comprehended by an intelligent layman. Excellent books on the pure mathematical theory of decision under uncertainty already exist; the present text is an endeavor to show how formal analysis of practical decision problems can be made to pay its way.

From the point of view taken in this book, there is no real difference between a "statistical" decision problem in which a part of the available evidence happens to come from a "sample" and a problem in which all the evidence is of a less formal nature. Both kinds of problems are analyzed by use of the same basic principles; and one of the resulting advantages is that it becomes possible to avoid having to assert that nothing useful can be said about a sample which contains an unknown amount of bias while at the same time having to admit that in most practical situations it is totally impossible to draw a sample which does not contain an unknown amount of bias. In the same way and for the same reason there is no real difference between a decision problem in which the long-run-average demand for some commodity is known with certainty and one in which it is not; and not the least of the advantages which result from recognizing this fact is that it becomes possible to analyze a problem of inventory control without having to pretend that a finite amount of experience can ever give anyone perfect knowledge of long-run-average demand. The author is quite ready to admit that in some situations it may be difficult for the businessman to assess the numerical probabilities and utilities which are required for the kind of analysis recommended in this book, but he is confident that the businessman who really tries to make a reasoned analysis of a difficult decision problem will find it far easier to do this than to make a *direct* determination of, say, the correct risk premium to add to the pure cost of capital or of the correct level at which to conduct a test of significance.

In sum, the author believes that the modern theories of utility and personal probability have at last made it possible to develop a really complete theory to guide the making of managerial decisions—a theory into which the traditional disciplines of statistics and economics under certainty and the collection of miscellaneous techniques taught under the name of operations research will all enter as constituent parts. He hopes, therefore, that the present book will be of interest and value not only to students and practitioners of inventory control, quality control, marketing research, and other specific business functions but also to students of business and businessmen who are interested in the basic principles of managerial economics and to students of economics who are interested in the theory of the firm. Even the teacher of a course in mathematical decision theory who wishes to include applications as well as complete-class and existence theory may find the book useful as a source of examples of the practical decision problems which do arise in the real world.

Because the purpose of this book is not to teach theory for its own sake but to show how theory can be applied to practical advantage in the real world, each new technique of analysis is applied to a realistic business problem as soon as it is introduced. Many of the most important principles are actually restated and reexplained in the contexts of several different kinds of decision problem, and for this kind of repetitiousness the author makes no apology. Learning depends on repetition; and if the rate of learning can be increased by printing up a few more sheets of white paper, the gain is well worth the cost. While some of the exposition could have been greatly condensed by the use of simple algebra and calculus, the author feels that even for students who have some familiarity with these techniques it is better to avoid their use in a statement of first principles. Justification of the steps in an argument by economic rather than purely formal reasoning develops an intuitive understanding of the essential features of a decision problem which is likely to be lost if attention is focused from the very first on problems of technical manipulation. On the other hand students who do have some command of mathematical technique may find it a useful exercise to supply proofs where these have been omitted from the text; and many of the examples and problems in the text can be easily modified to require the use of calculus rather than arithmetic for their solution. An appendix on gamma, beta, and related distributions has been added to the book to facilitate the assignment of problems of this sort.

The organization of the book reflects experience gained in teaching the subject in various ways to five successive classes. The basic concepts of decision theory—probability, expectation, and utility—are explained in three introductory chapters, and in the next five chapters (Part One of the book) these concepts are applied in a variety of situations where the required probability distributions can be easily assessed by direct reference to experience. It is only after the student has thus become reason ably familiar with the way in which probabilities are *used* that he is introduced in Part Two to some more powerful methods for the *computation* of probabilities—to the concepts of joint and conditional probability, the distributions associated with Bernoulli and Poisson processes, and the Normal distribution. After a foundation in *both* economic analysis and elementary probability theory has been laid, the student goes on in Part Three to face the special problem of the evaluation by means of Bayes' theorem of the information derived from a sample and to study some new distributions needed for this purpose. It is only after this subject has been thoroughly covered that the problem of deciding when it is economically advantageous to sample and when to stop sampling is taken up in Part Four. The four chapters which constitute Part Five of the book then explain the classical approach to the problems already analyzed from the Bayesian point of view in Parts Three and Four and show how the explicit introduction of losses into the classical analysis leads from operating characteristics to risk functions and how a reasoned comparison of risk functions over *all* values of the parameter under test then leads in the end to exactly the same results which were previously obtained by the explicit use of Bayes' theorem.

This division of the entire subject matter of the course into five separate major topics which are treated successively rather than simultaneously (as was necessary in earlier versions of this book which introduced sampling problems at the outset) has improved the rate at which the material can be absorbed to the point where the author is currently assigning nearly one chapter per 80-minute class session and teaches about three-fourths of the entire book in a one-semester course. The author's students, however, are the small fraction of second-year students at the Harvard Graduate School of Business Administration who voluntarily elect a course in decision theory which is well known to involve a very heavy work load, and obviously no such rate could be maintained with a less highly self-selected group of students. The author would guess that under ordinary conditions the book will prove to contain about the right amount of material for a full-year course, particularly if it is supplemented by some unstructured case problems or by mathematical lectures and exercises for students with a background of algebra and calculus.

The course can be shortened by omission of certain chapters which constitute excursions into interesting areas or problems of application but are not needed for the comprehension of later chapters; Chapters 15, 19, 20, 24, 32, and 36 are all of this sort. Even with these chapters omitted the student will have had a more than adequate foundation for a second course in statistics, e.g. in sampling theory or experimental design. A course covering only the basic principles of decision theory as such with an absolute minimum of attention to technicalities of probability theory can be given by using only Chapters 1 to 5, 7, 9, 10, 21, 22, 33, and 38. These 12 chapters explain every important basic principle discussed in the course, including the principles of optimal sample size and optimal sequential sampling, without the use of any mathematically derived probability distribution other than the binomial; they are an adequate preparation for the treatment of classical statistics in Chapters 39 to 42 if the examples and problems involving the use of the Normal distribution are omitted from those chapters. The other chapters in Parts One through Four of the course are there in part to develop the additional probability theory needed to handle a wider variety of applications and in part to develop special methods for the rapid analysis of a few of the most commonly occurring types of business decision problems—in particular, certain problems with linear losses. Without these methods which make it possible to obtain numerical answers to a fairly wide variety of examples in a reasonable amount of time, there is a real danger that the student will fail to gain any appreciation of the sensitivity or insensitivity of decisions and their associated losses to the various parameters of a decision problem; and without some appreciation of this sort the *practical* use of decision theory is very severely handicapped.

Exercises are provided at the end of each chapter. Most of them are intended to develop and test the student's comprehension of the theory expounded in the text, but some lead the student to extend this theory in some small degree. Completely worked solutions to all exercises of the latter sort and to about half those of the former sort will be found in the "Student's Manual" which accompanies the text. A slide rule is adequate computing equipment for the exercises with worked solutions, since in those problems the student needs only to verify that he understands how the computations were actually carried out, but the student who works problems on his own will usually find that a desk calculator will very greatly reduce the time required to arrive at a solution. These latter problems are well suited to work in a statistical laboratory.

The author's debt to his colleague Howard Raiffa and to his former colleague Arthur Schleifer, Jr., is far too great to describe adequately.

Mr. Schleifer assisted the author during the first 3 years of the development of the course represented by the present book. He read and made valuable criticisms of nearly every draft of every successive revision of every chapter, and he executed or directed the execution of all the computations and charting. Mr. Raiffa read the semifinal version of the manuscript with the most painstaking care and spent countless hours in showing an often stubborn author how it could and should be improved at many points by substituting logic for unsupported intuition. In particular, the three chapters of the Introduction were wholly recast as a result of these suggestions.

Mr. Gordon Kaufmann gave great assistance in preparing the Student's and Teacher's Manuals and corrected very many of the author's arithmetical lapses in text. The author was extremely fortunate to have Miss Alice Hynes (later Mrs. Paul O'Brien) as his secretary throughout the 4½ years during which the manuscript was being developed. Without her unusual skill it would have been quite impossible to make several rough drafts of each annual or semiannual revision of the notes and then to prepare stencils so that the latest version could be tested in the classroom.

Finally, the author would like to express his very deep gratitude to the administration of the Harvard Graduate School of Business Administration, both for substantially reducing his normal classroom assignments and for granting his every request for assistance at once and without question.

Robert Schlaifer

Contents

INTRODUCTION

The Problem of Decision under Uncertainty

CHAPTER 1

The Meaning of Probability

1.1 The Problem of Decision under Uncertainty

When all of the facts bearing on a business decision are accurately known—when the decision is made "under certainty"—careless thinking is the only reason why the decision should turn out, after the fact, to have been wrong. But when the relevant facts are not all known—when the decision is made "under uncertainty"—it is impossible to make sure that every decision will turn out to have been right in this same sense. Under uncertainty, the businessman is forced, in effect, to gamble. His previous actions have put him in a position where he *must* place bets, hoping that he will win but knowing that he may lose. Under such circumstances, a right decision consists in the choice of the best possible bet, whether it is won or lost after the fact. The following examples are typical of situations in which business decisions must be made and judged in this way.

An Inventory Problem. A retailer is about to place an order for a number of units of a perishable commodity which spoils if it is not sold by the end of the day on which it is stocked. Each unit costs the retailer $1; the retail price is $5. The retailer does not know what the demand for the item will be, but he must nevertheless decide on a definite number of units to stock.

A Scrap-allowance Problem. A manufacturer has contracted to deliver at least 100 good pieces of a nonstandard product at a fixed price for the lot. He feels virtually sure that there will be some defectives among the first 100 pieces produced; and since setting up for a second production run to fill out a shortage would cost a substantial amount of money, he wishes to schedule some additional pieces on the original run as a scrap allowance. On the other hand, once 100 good pieces have been produced the direct manufacturing cost of any additional production will be a total loss, and therefore he does not wish to make the scrap allowance excessively large. If the manufacturer knew exactly how many pieces would have to be produced in order to get exactly 100 good pieces, it would be easy to set the "right" size for the production order; but he must decide on some definite size for the order even though he does not know the "right" size.

An Investment Problem. A manufacturer is about to tool up for production of a newly developed product. This product can be manufactured by either of two processes, one of which requires a relatively small capital investment but high labor cost per unit produced while the other will have much lower labor costs but requires a much greater investment. The former process will thus be the better one if sales of the product are low while the latter will be better if sales are high; but the manufacturer must choose between the two processes without knowing what his sales will actually be.

A Marketing Problem. The brand manager for a certain grocery product is considering a change of package design in the hope that the new package will attract more attention on the shelf and thereby increase sales. He has done a certain amount of store testing and has found that during the test weeks sales of the new package were greater than sales of the old in some stores but that the contrary was true in other stores. He still feels uncertain whether adoption of the new package will increase or decrease his total national sales, but he must nevertheless either decide on one package or the other or else decide to spend more money on additional testing; in the latter case he must decide whether he should simply continue the test for a few more weeks in the same stores he has already used or spend still more money to draw new stores into his sample.

1.1.1 The Payoff Table

The essential characteristics of all four of these problems, and of all problems which we shall study in this course, are the following.

1. A choice must be made among several possible *acts*.
2. The chosen act will ultimately lead to some definite profit (possibly negative), but for at least some of the acts the amount of this profit is unknown because it will be determined by some *event* which cannot be predicted with certainty.

The first step in analyzing any such problem is to lay out all the possible acts and all their possible consequences in some systematic fashion, and we shall do this for the inventory problem as an example.

In the inventory problem, an "act" is a decision to stock some particular number of units; the "event" is the number of units which the customers will actually demand. If we suppose that the retailer's space limits the number of units stocked to a maximum of 5, then remembering that each unit stocked costs \$1 while each sale brings in \$5 of revenue we can describe the whole problem by a table like Table 1.1, where each column corresponds to a particular act while each row corresponds to a particular event. Such a table is known as a *payoff table.*

Table 1.1
Payoff Table for the Inventory Example

Event (number demanded)	Act (number of units stocked)					
	0	1	2	3	4	5
0	\$0	−\$1	−\$2	−\$3	−\$4	−\$5
1	0	+ 4	+ 3	+ 2	+ 1	0
2	0	+ 4	+ 8	+ 7	+ 6	+ 5
3	0	+ 4	+ 8	+12	+11	+10
4	0	+ 4	+ 8	+12	+16	+15
5 or more	0	+ 4	+ 8	+12	+16	+20

1.1.2 Comparison of Acts

If we compare any two acts (columns) in Table 1.1, we see that one of the two will be more profitable if certain events occur while the other will be more profitable if other events occur; but when we actually choose among these acts we are implicitly if not explicitly making a single, unconditional evaluation of each act. We are saying that in some sense one of the acts is "better" than any of the others. One conceivable way of evaluating the six possible acts of Table 1.1 is to look only at the worst possible result of each act and assign the value \$0 to the act "stock 0," the value −\$1 to the act "stock 1," and so forth, leading to the conclusion that "stock 0" is the best of all possible acts. Another conceivable way is to look only at the best possible result and assign the value \$0 to the act "stock 0," the value +\$4 to the act "stock 1," and so forth, leading to the conclusion that "stock 5" is the best of all possible acts.

Any sensible businessman will of course immediately reject all such simple but arbitrary procedures and will say that even though the retailer cannot predict demand with certainty he ought to know enough about his business and the product in question to have some convictions about what the demand is likely to be. If after weighing all the available information the retailer decides that there is very little chance that customers will demand less than 3 or more than 4 units, he will conclude that the only reasonable act is to stock either 3 or 4 units. Choice between these two acts will be a little more complex, since the larger stock will be only \$12 − \$11 = \$1 less profitable than the smaller if there is a demand for only 3 units while it will be all of \$16 − \$12 = \$4 more profitable if 4 units are demanded. Consequently the retailer will want to stock 4 units even if he believes that the chance of a demand for 4 is somewhat less than the chance of a demand for 3; it is only if he believes that the chance of a demand for 4 is relatively *very* slight that he will reduce his stock to 3 units.

Now this informal kind of reasoning works very well when the decision problem is relatively simple, but one quickly becomes confused when the problem is even slightly more complex. Even in our very simple example, it will be hard for the retailer to see through to a satisfying conclusion if he thinks that there is a substantial chance that demand may have any of three or four different values, and in larger problems of the same sort he may well consider a hundred or a thousand different values as possible. What we would like to do, therefore, is find some way of *systematizing* the kind of analysis which a reasonable man uses in simple problems so that it can be effectively applied in more complex problems.

If we look back at the reasoning used by our hypothetical retailer, we see that in essence he proceeded in two steps: he first gave a numerical *value* to the consequence of each possible act given each possible event, but he then attached more *weight* to the consequences corresponding to certain events (demand 3 or 4) than he did to the others. This suggests that it may be possible to systematize the reasoning underlying *any* decision under uncertainty by proceeding as follows:

1. Attach a definite numerical *value* to the consequence of every possible act given every possible event.
2. Attach a definite numerical *weight* to every possible event.
3. For each act separately, use these weights to compute a *weighted average* of all the values attached to that act.
4. Select the act whose weighted-average value is highest.

Our hope is that we can find rules for using the businessman's own knowledge and beliefs in carrying out steps 1 and 2 in such a way that he will *want* to choose the act with the highest computed value instead of relying on mere inspection of a mass of numbers and informal reasoning of the kind described above. If we are to have confidence in these rules in complex situations, they must yield values which seem reasonable to us when applied in very simple situations, and for this reason many of the examples which we shall use in developing these rules will be artificial ones which avoid the complexities of practical business decisions in order to present their really essential features in the simplest possible form. Because the heart of the problem is the uncertainty concerning the event, we shall begin by developing the rules for attaching weights to events.

1.2 Events

Before we even start to assign numerical weights to a set of events some one of which will determine the consequence of any act we choose, we obviously must have in mind a clear and complete description of the events which may occur. We usually have considerable latitude in defining the possible events in a given problem, but certain rules must be followed if we are to avoid hopeless confusion.

1.2.1 Collectively Exhaustive Events

If before we started to analyze the inventory problem of Table 1.1 the retailer had told us that he was absolutely convinced that there would be a demand for at least 2 units, we could just as well have simplified Table 1.1 by eliminating the rows describing the consequences of the events "demand 0" and "demand 1." In general, impossible events may be totally disregarded if it is convenient to do so, and it is to be emphasized that there is no need to "prove" that an event is impossible before it is eliminated. Our object is to arrive at results which the businessman *wants to accept*, and therefore an event is impossible for our purposes whenever the businessman wants to treat it as impossible.

It is obvious, on the other hand, that we must keep *all* the *possible* events in mind in analyzing any decision problem, since if we fail to include some of the possible events in the payoff table the corresponding consequences will not be duly considered in evaluating the various acts. The same thing can be stated the other way around: the basic list of events must be complete in the sense that *some one of the events on the list is bound to occur*. The events on such a list are called *collectively exhaustive*.

1.2.2 Mutually Exclusive Events

In the inventory example of Table 1.1, demand for each specific number of units from 0 to 4 inclusive was treated as a separate event but demands for all numbers of units above 4 were treated as constituting the same event "demand for 5 or more." Obviously we *could* have treated a demand for exactly 5 units as a separate event and assigned it a separate line in Table 1.1, and similarly for any larger number of units, but nothing was to be gained by so doing because for every act under consideration the consequences of the event "demand for 5" were identical to the consequences of the event "demand for 6" or the event "demand for 7" and so forth.

Careless grouping of events can easily lead to confusion, however. It is obvious that potentially separate events must not be grouped if their consequences differ for any act under consideration. We cannot treat "demand for 3 or 4" or "demand for 4 or more" as a single event in constructing a payoff table for our inventory example. What is often less obvious is that we must not have events with *overlapping definitions* on our list even if it is possible to give a clear description of the consequences of all acts in terms of such a list.

Suppose, for example, that we are given a choice of one or the other of two tickets in a lottery to be conducted by drawing one ball from an urn containing four kinds of balls: dotted red, striped red, dotted green,