AD-A016 276

MODELING AND DECISION ANALYSIS

Steven Nobumasa Tani

Stanford University

Prepared for:

Advanced Research Projects Agency
National Science Foundation
Office of Naval Research

30 June 1975

ADA016276

Research Report No. EES-DA-75-3
June 30, 1975

304177

MODELING AND DECISION ANALYSIS

STEVEN NOBUMASA TANI

DECISION ANALYSIS PROGRAM

Professor Ronald A. Howard
Principal Investigator

DEPARTMENT OF ENGINEERING-ECONOMIC SYSTEMS

Stanford University
Stanford, California 94305

SPONSORSHIPS

Advanced Research Projects Agency, Human Resources Research Office, ARPA Order No. 2449, monitored by Engineering Psychology Programs, Office of Naval Research, Contract No. N00014-67-A-0112-0077 (NR197-024).

National Science Foundation, NSF Grant GK-36491.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: EES-DA-75-3	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): "Modeling and Decision Analysis"	5. TYPE OF REPORT & PERIOD COVERED: Technical 5/1/73 to 5/31/75
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Steven Nobumasa Tani	8. CONTRACT OR GRANT NUMBER(s): N00014-67-A-0112-0077
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The Board of Trustees of the Leland Stanford Junior University, c/o Office of Research Administrator, Encina Hall, Stanford, California 94305	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 000000 ARPA Order #2449
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency Human Resources Research Office 1400 Wilson Blvd., Arlington, Virginia 22209	12. REPORT DATE: June 30, 1975
	13. NUMBER OF PAGES: 149 pages
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Engineering Psychology Programs, Code 455 Office of Naval Research 800 No. Quincy Street, Arlington, Virginia 22217	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

A dissertation submitted to the Department of Engineering-Economic Systems and the Committee on Graduate Studies of Stanford University in partial fulfillment of the Requirements for the Degree of Doctor of Philosophy, June 1975.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

DECISION ANALYSIS — DECISION THEORY
DECISION-MAKING — MODELING

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We use mathematical modeling in decision analysis to help us obtain a "better" profit lottery than we can assess directly. The concept of the authenticity of probabilities is introduced to define the measure of "goodness" of the profit lottery. The role of modeling is to simplify our assessment task through the decomposition of the profit lottery. However, budgetary constraints force us to make approximations in the modeling process and thereby cause us to misstate the profit lottery. (Continued on reverse side)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 (continued)

The models used in a decision analysis should be regarded as subjective expressions of our uncertainty rather than as objective descriptions of the real-world.

A methodology is presented that quantitatively relates the modeling approximations made in a decision analysis to the results of the analysis.

Abstract

We use mathematical modeling in decision analysis to help us obtain a "better" profit lottery than we can assess directly. The concept of the authenticity of probabilities is introduced to define the measure of "goodness" of the profit lottery. The role of modeling is to simplify our assessment task through the decomposition of the profit lottery. However, budgetary constraints force us to make approximations in the modeling process and thereby cause us to misstate the profit lottery. The models used in a decision analysis should be regarded as subjective expressions of our uncertainty rather than as objective descriptions of the real-world.

A methodology is presented that quantitatively relates the modeling approximations made in a decision analysis to the results of the analysis.

Acknowledgements

I wish to express my profound gratitude to Professor Ronald A. Howard, my teacher and advisor, for his invaluable guidance and patient encouragement throughout the course of this research.

Thanks also to Professor Richard D. Smallwood, who first sparked my interest in this topic and with whom I have debated long and fruitfully about modeling, and Professor Donald A. Dunn for serving on my reading committee.

I wish to thank my colleagues Peter Morris, Stephen Barrager and Bruce Judd for reading early drafts of this thesis and offering helpful comments and suggestions.

I am grateful to the faculty and staff of the Department of Engineering-Economic Systems for making my long stay there a happy one. Special thanks go to Mrs. Doris Troiani, whose unerring guidance in administrative matters and whose moral support helped me through many years of study.

This research was sponsored in part by the Advanced Research Projects Agency under Contract Number N00014-67-A-0012-0077.

Finally, I dedicate this work to the memory of my father, Henry N. Tani, and to my mother, Rose S. Tani. All that I am I owe to them and I can only hope that my life honors theirs.

Table of Contents

Table of Contents, continued

List of Tables

List of Figures

INTRODUCTION

To the Wizard and to the Clairvoyant, decision making is a simple matter. The omnipotent Wizard has the power to change anything in the world to suit his own taste. To him, there is no such thing as a decision problem; he merely causes to happen whatever he wishes. The omniscient Clairvoyant, although lacking the Wizard's power to change events, possesses the power to foretell the future perfectly. To him, there is no uncertainty in decision making; he merely chooses the course of action whose consequences he most desires.

By contrast, we lowly mortals cannot pretend to possess the Wizard's or the Clairvoyant's powers in decision making. But there is a superbeing whom we do try to emulate -- the Elicitor. The Elicitor can neither change nor foretell events and he is therefore uncertain about the future. However, he possesses the special ability to fully and accurately express his uncertainty in the form of probability statements. In a decision problem, the Elicitor assesses his probabilities on the future consequences of the alternative courses of action and, using his utility function, calculates the expected utility for each alternative. He then chooses the course of action having the highest expected utility. He is not assured that his decision will lead to the most desired outcome, but he is confident that his actions are wholly consistent with his preferences and with his uncertain understanding of the future.

In decision analysis, seeking to imitate the Elicitor, we likewise make probability statements about the future consequences of our actions and calculate the resulting expected utility for each alternative. But we unfortunately do not possess the Elicitor's ability to directly assess probabilities that fully and accurately express our uncertainty. Recognizing this, we use modeling to help us obtain the requisite probability statements.

Modeling is the source of most of our dissatisfaction with the results of particular decision analyses. While we do not contest the validity of decision theory, we often complain that the models used in a decision analysis are "too simplistic" or "not realistic enough" or "not believable" and we therefore regard the results of the decision analysis with doubting eyes.

How should we deal with our dissatisfaction about modeling in decision analysis? What do we mean by "goodness" in a model and can we quantify it? Can we, for instance, define an index of "realism" or of "credibility" on models? How should we choose among alternative models? For example, can we use the notion of a probability-space of models? And how should we decide when to do more modeling? For instance, could we use the concept of the value of perfect modeling as an analogue to the value of perfect information?

In this dissertation, I offer a way to think about modeling in decision analysis so that we can deal with our dissatisfactions meaningfully. Chapter 1 provides a philosophical perspective on the role of modeling in decision analysis. Chapter 2 presents a methodology with which we can quantify our dissatisfaction about the modeling in a decision analysis and relate it to the results of the analysis. Chapter 3 is an example illustrating the use of the methodology and Chapter 4 is an extension of the methodology to stochastic models.

CHAPTER 1

A PERSPECTIVE ON MODELING IN DECISION ANALYSIS

1.1 Introduction

In theory, we do not need to use mathematical modeling to perform decision analysis. Analytically, a decision problem is defined by two variables: a decision variable d, representing the alternative courses of action open to us, and an outcome variable v, representing the relevant consequences of our actions. In decision analysis [7,8], we need only assess for each value of d the conditional probability distribution on v given d, $\{v|d,\delta\}$, called the profit lottery and state our risk preference by specifying a utility function on v, $u(v)$. Then, we can calculate the expected utility of each profit lottery:

$$\bar{u}_d = \int_v \{v|d,\delta\} \; u(v)$$

and identify that value d^* of the decision variable corresponding to the profit lottery with the highest expected utility (i.e., the most preferred profit lottery). The optimal alternative is the one specified by d^*.

Although superfluous in theory, modeling is nevertheless indispensable to decision analysis in practice. In theory, we assume that we can directly assess the profit lottery, but in practice, we are unable to do so satisfactorily. Consequently, we turn to modeling to help us obtain the profit lottery.

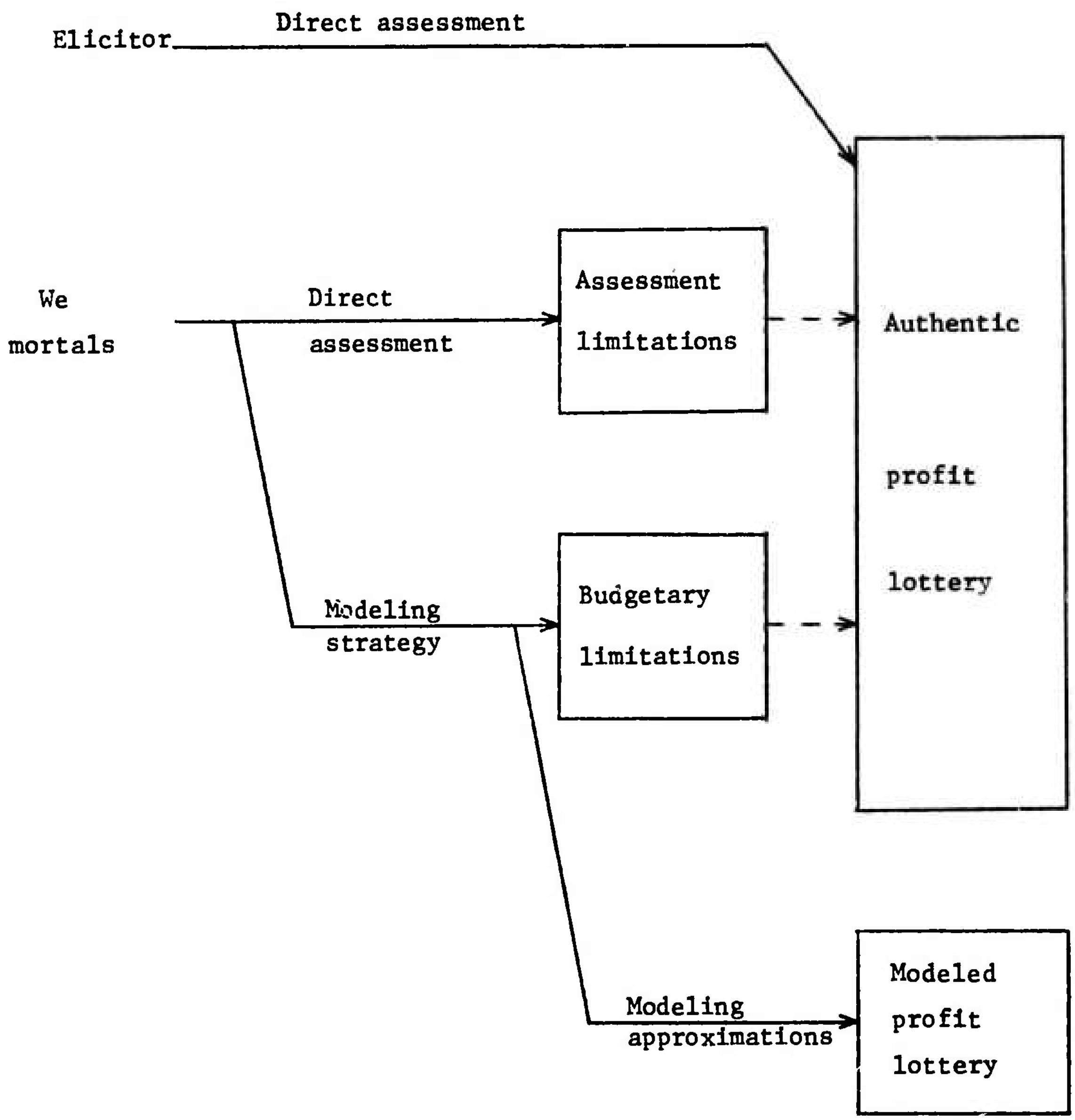
Our ideal in decision analysis is not to construct the perfect model, but rather to obtain the authentic profit lottery -- the one that accurately expresses our uncertainty about the future. (See Figure 1.1 for a conceptual "roadmap".) We do not have the Elicitor's ability to assess the authentic profit lottery directly, so we employ the modeling strategy, which would yield the authentic profit lottery if successfully executed. However, budgetary constraints force us to make approximations in the modeling process; consequently, the profit lottery that we obtain through modeling is not the authentic profit lottery. The modeling approximations, then, are the sole source of our dissatisfaction.

1.2 Authentic Probabilities

We use modeling in decision analysis because we do not believe that the directly assessed profit lottery is good enough.

But what is "goodness" in a profit lottery? The profit lottery, as a subjective probability statement [3,5,16], is the quantified expression of our beliefs about the likelihood of occurrence of real-world events beyond our immediate perception. A good probability statement is simply one that accurately and fully expresses our beliefs. To denote a good probability statement, I use the term "authentic", which the dictionary defines as "worthy of acceptance and belief."

Figure 1.1: A Conceptual "Roadmap"



Authenticity is not the same as trueness. The only test of goodness of a probability statement is whether or not we believe it, not whether it is true or false. Indeed, it is not even meaningful to speak of a probability statement as being true or false.

Compare, for example, the two statements: "It will rain tomorrow." and "The probability of rain tomorrow is 3/4." The first statement is an assertion about the real-world and is either true or false. We can determine its goodness (i.e., its trueness) by observing the world (tomorrow). The second statement, on the other hand, does not describe the real-world, but instead expresses our uncertainty about the world. It makes no sense to say that it is true or false, but only that it is authentic or inauthentic. We cannot ascertain its goodness (i.e., its authenticity) by appealing to the real-world, but rather only to our beliefs.

The authenticity of a probability statement depends on our state of information. A probability statement that we accept as authentic based on one set of information may become inauthentic when we receive additional information. For example, if we observe rainclouds gathering overhead, our authentic probability for rain tomorrow may change.

But how do we obtain authentic probabilities? Clearly, the authentic probability for an event must reflect all of the information we possess relevant to that event. Therefore, to assess the authentic probability, we must review our entire state of information pertaining to the event and judge the likelihood of its occurrence on the basis of that information.

This is a formidable task for almost any event. Since we believe that the world is highly interdependent, our state of information is an extensive and complicated web of knowledge. Hence, we must retrieve, organize and process a large amount of information to assess an authentic probability.

Consider, for example, the probability that the winner of the 1980 U.S. Presidential election is a member of the Republican Party. To assess the authentic probability for this event, we must consider every set of circumstances that might lead to its occurrence. We must review our knowledge of historical trends, of the current state of the Republican Party and of the intentions and qualifications of potential candidates. We must relate the authentic probability to our uncertainty about intervening events, such as the outcome of the 1976 election and the possibility of war or depression before 1980.

We encounter two major difficulties in trying to assess the authentic probability. First, we find that the task of considering everything that might affect the outcome of the event is seemingly endless. And second, we find it virtually impossible to perform the requisite processing of the information without external computational assistance.

Let us imagine the existence of a person, called the Probabilist, who is capable of performing upon request any calculation using the rules of probability calculus (e.g., Bayes' Theorem, expansion, change of variable). Then, we can state the following operational definition of authenticity: The authentic probability for an event is the one we would obtain if we could spend an unlimited amount of time in introspection and if we had available the services of the Probabilist.

It should be readily apparent that the authentic probability is, in most cases, an unattainable ideal. Even after lengthy introspection, we can almost always think of something relevant that we have not yet considered. And, of course, we do not possess the computational capabilities of the Probabilist.

1.3 Operative Probabilities

Since we cannot obtain authentic probabilities, we must use the probabilities that are based on only partial consideration of our state of information. These probabilities I call "operative".

As a matter of notation, we can represent our incomplete consideration of the state of information $\mathcal{S}$ as $C(\mathcal{S})$. Then, our authentic and operative probabilities for event E are:

$$\{E|\mathcal{S}\} = p_a \quad \text{Authentic probability}$$

$$\{E|C(\mathcal{S})\} = p_o \quad \text{Operative probability}$$

It is useful here to distinguish between two kinds of uncertainty. The first kind, called primary uncertainty, is the uncertainty due solely to the finiteness of our state of information and is expressed by authentic probabilities. The second kind, called secondary uncertainty, is the additional uncertainty due to the incomplete consideration of our state of information. Whereas primary uncertainty can be resolved only by receiving more information, secondary uncertainty can be resolved by further introspection and calculation.

Conceptually, we can represent our secondary uncertainty as a conditional probability distribution on the authentic probability given partial consideration of our state of information:

$$\{p_a | C(\delta)\}$$

Then, the mean of this distribution is the operative probability:

$$p_o = \langle p_a | C(\delta) \rangle$$

and the variance ${}^v\langle p_a | C(\delta) \rangle$ is a measure of how much secondary uncertainty remains to be resolved and thus how "close" the operative probability is to the authentic probability.

As we spend more time in introspection and calculation, we reduce the amount of secondary uncertainty and thereby make the operative probability converge to the authentic probability.

For example, the operative probability of a Republican victory in the 1980 election that we assign after only a moment's thought is "far" from the authentic probability; that is, since much secondary uncertainty remains, the operative probability may change significantly with further introspection and calculation. On the other hand, the operative probability that we assign to getting a head on the next flip of a coin is very "close" to the authentic probability because little secondary uncertainty remains and further consideration of our state of information is unlikely to markedly change the operative probability; in this case, we can say that the operative probability is virtually authentic.

But what difference does it make how "close" the operative probability is to the authentic probability? If we must act without further consideration of our state of information, then it makes no difference how "close" the operative probability is to the authentic probability. However, if we are able to further consider our state of information before acting, then the "farther" the operative probability is from the authentic probability, the greater the value of doing some more introspection and calculation.

For example, suppose that we are offered a choice between a sure \$60 and a lottery yielding \$100 if a specified event occurs and nothing otherwise; the event may be either a Republican victory in the 1980 election or a head on the next flip of a coin. Suppose further that our operative probability p_o of a 1980 Republican victory is identical to our operative probability q_o of a head on the next flip of the coin, but that p_o is "far" from the authentic probability while q_o is virtually authentic:

$$p_o = q_o = \tfrac{1}{2}$$

$$\{p_a | C(\delta)\} = \begin{cases} 1 & 0 \leq p_a \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\{q_a | C(\delta)\} = \delta(q_a - \tfrac{1}{2})$$

Then, if we are prohibited from doing further introspection or calculation, it does not matter which event is used in the lottery; in either case, the expected value of the lottery is $50 and our decision is to choose the sure $60. However, if we are permitted the opportunity to introspect and calculate before choosing, it does matter which event is used in the lottery. If the lottery depends on the coin flip, no amount of introspection or calculation will change the expected value of the lottery, so we would always take the sure $60. In this case, further introspection and calculation has no value. On the other hand, it does have value if the lottery depends on the 1980 election. In this case, further introspection and calculation may reveal that the authentic probability is such that we would choose the lottery instead of the sure $60 ($p_a$ >.6). The expected value of the decision given complete introspection and calculation is:

$$\bar{v} = \int_{p_a} \langle v|p_a \rangle \{p_a|C(\mathcal{S})\} \; dp_a$$

$$= \int_0^{.6} 60 \; dp_a \quad + \quad \int_{.6}^{1} 100 \; p_a \; dp_a$$

$$= \quad 36 \qquad\qquad + \quad 32$$

$$= \quad \$68$$

Therefore, the expected value of complete introspection and calculation (i.e., the value of obtaining the authentic probability) is $68 - $60 = $8.

The value of further consideration of our state of information is analogous to the value of receiving additional information. In both cases, the value depends on how much uncertainty remains to be resolved and on how much is at stake in the decision.

1.4 The Modeling Strategy

In decision analysis, then, we want the authentic profit lottery -- the one and only probability distribution that accurately expresses our beliefs about the future consequences of our actions.

But in most decision situations, our state of information about the decision environment is much too complicated for our simple human minds to consider completely.

For example, suppose that our decision is whether or not to manufacture and market a new product. We might be able to identify scores of factors that affect the outcome of our decision, including the behavior of our competitors, the general economic conditions, the effectiveness of our advertising, the costs of raw materials, the efficiency of the manufacturing process and the regulatory behavior of the government. The authentic profit lottery must express our uncertain understanding of how each of these factors affects our profit and how each will behave in the future.

Clearly, we would be hard pressed to process all of this information mentally. Consequently, the profit lottery that we assess directly is very "far" from the authentic profit lottery and the value of resolving our secondary uncertainty is high.

Recognizing this, we turn to modeling to help us resolve our secondary uncertainty. In essence, the modeling strategy is one of divide-and-conquer; in modeling, we decompose the profit lottery into smaller pieces that our simple minds can handle comfortably. Then, we rely on external means of calculation to reassemble the pieces to obtain the profit lottery.

The modeling strategy is as follows:

1. Identify a set of real-world factors on which profit is believed to depend, representing them as state variables, denoted by the vector $\underline{s}$.

2. Encode our uncertain understanding of the dependence relationship between profit v and the decision and state variables d and $\underline{s}$ as a conditional probability distribution $\{v|d,\underline{s},\mathcal{S}\}$.

3. Encode our uncertainty about the future behavior of the state variables as a probability distribution $\{\underline{s}|\mathcal{S}\}$.

4. Using external means of calculation, determine the profit lottery via the expansion equation:

$$\{v|d,\mathcal{S}\} = \int_{\underline{s}} \{v|d,\underline{s},\mathcal{S}\}\{\underline{s}|\mathcal{S}\}$$

Note that the modeling strategy does not relieve us of the necessity to assess probabilities. Rather, it merely substitutes for the direct assessment of the profit lottery the assessment of probabilities on the state variables and on the dependence relationship.

The key to success of the modeling strategy, then, is selecting a set of state variables such that we can handle the resulting assessment task satisfactorily.

If we can assess virtually authentic probabilities on the state variables and on the dependence relationship, then the modeling strategy would yield the authentic profit lottery, since the expansion equation is tautologically true.

1.5 The Abridged Modeling Strategy

Unfortunately, we can seldom fully execute the modeling strategy because we are limited by economic constraints on the decision analysis itself. If a decision analysis is to be worthwhile, its cost must be small relative to the resources allocated in the decision. Therefore, in any decision analysis, we must strictly limit the time and effort devoted to obtaining the profit lottery.

Full execution of the modeling strategy, by contrast, would require much time and effort. Generally, the assessment of virtually authentic probability distributions on the state variables and on the dependence relationship would be very time-consuming. Furthermore, because we cannot analytically perform integration over an arbitrary continuous function, as required by the strategy, we would need to perform the integration numerically, which would also be exceedingly expensive.

Recognizing that we are unable to fully execute the modeling strategy within the budgetary constraints, we abbreviate it. The abridged modeling strategy is as follows:

1. Identify a set of real-world factors on which profit is believed to depend, representing them as state variables, denoted by the vector $\underline{s}$.

2. Identify a deterministic function g that approximates the dependence relationship between profit v and the decision and state variables d and $\underline{s}$:

$$v \approx g(d,\underline{s})$$

The function g is commonly called the "model".

3. Approximately encode our uncertainty about the future behavior of the state variables as a discrete probability function $p(\underline{s})$.

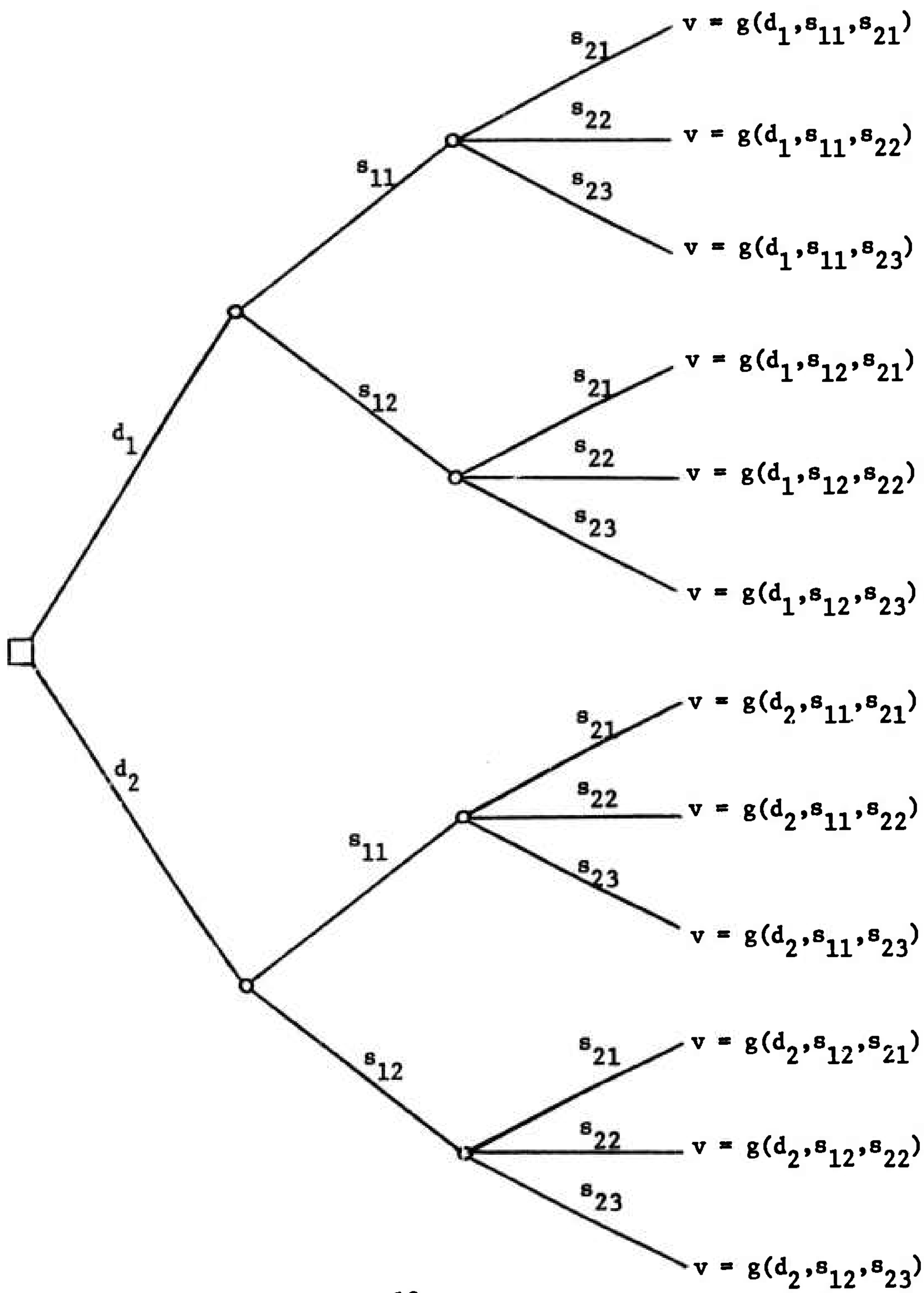
4. Using external means of computation, calculate the profit lottery via the expansion equation:

$$\{v|d,\mathcal{S}\}_m = \sum_{\underline{s}} \delta(v - g(d,\underline{s}))\ p(\underline{s})$$

The abbreviation of the modeling strategy results in the familiar decision tree representation of the decision problem. [23] (See Figure 1.2) The discrete steps of the state variables are represented as "branches" at the chance nodes of the tree and the profit assigned to each "tip" of the tree is given by the deterministic function:

$$v = g(d,\underline{s})$$

Figure 1.2: Decision Tree Representation



The degree to which we must restrict the number of state variables and the level of discretization of their distributions depends, of course, on the budgetary constraints of the particular decision analysis. Usually, the restriction is quite severe because the size of the decision tree grows geometrically with the number of variables and with the number of discrete steps in the distributions. More specifically, if we have N state variables and if we discretize the i-th state variable into m_i steps, the number of tips k in the decision tree for each decision alternative is given by:

$$k = \prod_{i=1}^{N} m_i$$

For each of the k tips, we must evaluate the function $v = g(d,\underline{s})$ to determine the profit associated with that tip and we must multiply together N probabilities to find the probability corresponding to that value of profit.

Clearly, the amount of computation required to solve a decision tree can easily become prohibitively large. For example, if we have ten state variables and if each is discretized into five steps, our decision tree will have nearly ten million tips for each decision alternative. Solving this tree would tax the capabilities of almost any computer and would incur an exceedingly high cost.

Consequently, to meet the budgetary constraints, we deliberately ignore our uncertainty about some of the state variables in the interest of computational economy. We select from the state variables only those very few that have the greatest effect on profit, as revealed by sensitivity analysis, and designate them as aleatory variables. [19,27] We assess the probability distribution on each aleatory variable and discretize it into only a few steps. We designate the remaining state variables as non-aleatory and, disregarding our uncertainty about them, fix each of them at a single representative value. In effect, we discretize each non-aleatory variable into just one step instead of into several. Typically, we have fewer than ten aleatory variables and we rarely discretize a distribution into more than three steps.

The profit lottery that we obtain through the abridged modeling strategy is not the authentic profit lottery. The profit lottery from the abridged modeling strategy is:

$$\{v|d,\mathcal{S}\}_m = \sum_{\underline{s}} \delta(v - g(d,\underline{s}))\, p(\underline{s})$$

From the full modeling strategy, the authentic profit lottery is:

$$\{v|d,\mathcal{S}\} = \int_{\underline{s}} \{v|d,\underline{s},\mathcal{S}\}\ \{\underline{s}|\mathcal{S}\}$$

The discrepancy between the modeled and authentic profit lotteries is due solely to the two approximations made in the abridged modeling strategy:

1. $\{v|d,\underline{s},\mathcal{S}\} \approx \delta(v - g(d,\underline{s}))$

2. $\{\underline{s}|\mathcal{S}\} \approx p(\underline{s})$

I call the first the Dependence Approximation and the second the Distribution Approximation. In Chapter 2, I present a methodology for estimating the size of the discrepancy in the profit lotteries caused by these two modeling approximations.

1.6 The Subjective Nature of the Model

In the abridged modeling strategy, we attribute all of our uncertainty about profit to our uncertainty about the state variables and none to uncertainty about the dependence relationship between profit and the state variables.

This arrangement, although motivated by budgetary considerations, is also conceptually appealing because it mirrors the way in which we visualize the world. According to our Western world-view, we believe that phenomena do not occur spontaneously, but are caused or influenced by other phenomena; hence, we "explain" the occurrence of a phenomenon by establishing the occurrence of its antecedents. Furthermore, we believe that the interactions among the phenomena are goverened by permanent rules of relationship (i.e., natural "laws").

It is the role of science to increase our understanding of the world by discovering and codifying these rules of relationship through the careful observation of the universe.

Applying this world-view to the decision situation, we believe that a particular level of profit does not simply happen, but is caused by certain extraneous factors. We believe, then, that there exists a set of state variables $\underline{S}$ that completely determine profit through a fixed real-world relationship W :

$$v = W(d,\underline{S})$$

I call such a set of state variables "complete".

We believe that, if we could predict the future behavior of the state variables $\underline{S}$ and if we knew W , we could then predict profit exactly. We are thus led to idealize the modeling process. In the idealized modeling scheme, we determine W by examining objective real-world data and encode our uncertainty about the future behavior of the state variables as a probability distribution $\{\underline{S}|\mathcal{S}\}$. Then, through the relationship W , this uncertainty about the state variables translates directly into uncertainty about profit, expressed as the profit lottery:

$$\{v|d,\mathcal{S}\} = \int_{\underline{S}} \delta(v - W(d,\underline{S})) \; \{\underline{S}|\mathcal{S}\}$$

The abridged modeling strategy mimics this idealized modeling scheme in that no uncertainty is attributed to the dependence relationship between profit and the state variables. Rather, all uncertainty about profit is attributed solely to uncertainty about the state variables $\underline{s}$ acting through the deterministic model g :

$$\{v|d,\mathcal{S}\}_m = \sum_{\underline{s}} \delta(v - g(d,\underline{s}))\ p(\underline{s})$$

But this arrangement is a misrepresentation of our beliefs. The modeling that we actually perform in a decision analysis falls short of the idealized modeling scheme in two major respects. First, the set of state variables $\underline{s}$ that we specify for use in the modeling strategy is generally not complete. Typically, in a complex decision problem, although we can identify many different factors that influence profit, we select as state variables only those few that have the greatest effect on profit in order to restrict the assessment and computational costs. Consequently, we believe that profit depends not only on these specified state variables but on other, unspecified, variables as well. Even if we knew the value of each specified variable, we would remain uncertain to some extent about profit because of the effects of the unspecified variables.

Let us represent the unspecified variables as the vector $\underline{z}$. Then $\underline{s}$ and $\underline{z}$ together constitute a complete set of state variables:

$$(\underline{s},\underline{z}) = \underline{S}$$

Assume for the moment that we know W , the real-world relationship between profit and this complete set of state variables, $v = W(d,\underline{s},\underline{z})$.

Then, our uncertainty about the dependence relationship between profit v and the specified state variables $\underline{s}$ is:

$$\{v|d,\underline{s},\mathcal{s}\} = \int_{\underline{z}} \{v|d,\underline{s},\underline{z},\mathcal{s}\}\ \{\underline{z}|\underline{s},\mathcal{s}\}$$

$$= \int_{\underline{z}} \delta(v - W(d,\underline{s},\underline{z}))\ \{\underline{z}|\underline{s},\mathcal{s}\}$$

However, we generally do not possess sufficient real-world data to determine W ; this is the second respect in which we fall short of the idealized modeling scheme. Although we believe that there does exist a fixed real-world relationship between profit and the complete set of state variables, we are not knowledgeable enough to say what it is.

Let W_i be one of several alternative relationships. Then, letting $\{W_i|\mathcal{s}\}$ be the probability that W_i is the actual relationship, our uncertainty due to this insufficiency of data can be represented as:

$$\{v|d,\underline{s},\underline{z},\mathcal{s}\} = \int_{i} \delta(v - W_i(d,\underline{s},\underline{z}))\{W_i|\mathcal{s}\}$$

We see, then, that our uncertainty about the dependence relationship between profit v and the specified state variables $\underline{s}$ has two sources -- our uncertainty about the effects of the unspecified variables $\underline{z}$ and our uncertainty about W_i due to a lack of sufficient data:

$$\{v|d,\underline{s},\mathcal{S}\} = \int_{\underline{z}} \int_{i} \delta(v - W_i(d,\underline{s},\underline{z}))\ \{W_i|\mathcal{S}\}\ \{\underline{z}|\underline{s},\mathcal{S}\}$$

Clearly, we ignore both of these sources of uncertainty when we assert in the abridged modeling strategy that profit is completely determined by the specified state variables $\underline{s}$ acting through the model g :

$$v = g(d,\underline{s})$$

That is, in making this assertion, we assume clairvoyance both on $\underline{z}$ given $\underline{s}$ and on W_i :

$$\{\underline{z}|\underline{s},\mathcal{S}\} = \delta(\underline{z} - \underline{z}_0(\underline{s})) \quad \text{for some fixed } \underline{z}_0(\underline{s})$$

and

$$\{W_i|\mathcal{S}\} = \delta(W_i - W_k) \quad \text{for some fixed } W_k$$

Assuming this clairvoyance, we would have:

$$\begin{aligned}
\{v|d,\underline{s},\mathcal{S}\} &= \int_{\underline{z}} \int_{i} \delta(v - W_i(d,\underline{s},\underline{z}))\{W_i|\mathcal{S}\}\ \{\underline{z}|\underline{s},\mathcal{S}\} \\
&= \int_{\underline{z}} \int_{i} \delta(v - W_i(d,\underline{s},\underline{z}))\ \delta(W_i - W_k)\ \delta(\underline{z} - \underline{z}_0(\underline{s})) \\
&= \delta(v - W_k(d,\underline{s},\underline{z}_0(\underline{s}))) \\
&= \delta(v - g(d,\underline{s}))
\end{aligned}$$

$$\text{where } g(d,\underline{s}) = W_k(d,\underline{s},\underline{z}_0(\underline{s}))$$

But we do not possess such clairvoyance; we are indeed uncertain about the dependence relationship between profit and the specified state variables. Therefore, we should regard the approximation made in the abridged modeling strategy:

$$\{v|d,\underline{s},\mathcal{S}\} \simeq \delta(v - g(d,\underline{s}))$$

as just that -- an approximation made for budgetary reasons. Despite the similarity in appearances, we must not confuse the role of the model g in the abridged modeling strategy with that of the relationship W in the idealized modeling scheme. Unlike W , the model g does not represent a real-world relationship between profit and the state variables, but is only a convenient surrogate for the probabilistic relationship that fully expresses our uncertainty. The model g is not an objective description but rather a subjective expression.

Several writers [14,17,26] have suggested a probability-space of models as a vehicle for expressing our uncertainty about the dependence relationship. The use of a model space is attractive because it allows us to employ Bayesian techniques to choose among several alternative model forms in the face of real-world data.

If we use such a vehicle, we must be careful how we interpret the probability assigned to each model in the space. As we have seen, unless the specified set of state variables is complete, there can be no "correct" model g that deterministically relates profit to the state variables; therefore, it is meaningless to speak of the probability that a particular model is the "correct" one or that it embodies the actual relationship.

However, there is a way that we can make a meaningful interpretation of the probability-space of models. Suppose that we are uncertain about which of several relationships W_i is the actual relationship between profit and the complete set of state variables $(\underline{s},\underline{z})$:

$$\{v|d,\underline{s},\underline{z},\delta\} = \int_i \delta(v - W_i(d,\underline{s},\underline{z})\{W_i|\delta\}$$

Then, we can write the conditional mean of v given d and $\underline{s}$ as:

$$\begin{aligned}
\langle v|d,\underline{s},\delta\rangle &= \int_v v\{v|d,\underline{s},\delta\} \\
&= \int_v v \int_{\underline{z}} \int_i \delta(v - W_i(d,\underline{s},\underline{z}))\{W_i|\delta\}\{\underline{z}|\underline{s},\delta\} \\
&= \int_i \{W_i|\delta\} \int_{\underline{z}} \int_v v\ \delta(v - W_i(d,\underline{s},\underline{z}))\{\underline{z}|\underline{s},\delta\} \\
&= \int_i \{W_i|\delta\} \int_{\underline{z}} W_i(d,\underline{s},\underline{z})\{\underline{z}|\underline{s},\delta\} \\
&= \int_i \{W_i|\delta\}\ g_i(d,\underline{s})
\end{aligned}$$

where we have defined:

$$\begin{aligned}
g_i(d,\underline{s}) &= \int_{\underline{z}} W_i(d,\underline{s},\underline{z})\{\underline{z}|\underline{s},\delta\} \\
&= \langle v|d,\underline{s},W_i,\delta\rangle
\end{aligned}$$

Thus, we can expand the conditional mean $\langle v|d,\underline{s},\delta\rangle$ over a space of models, where the probability assigned to model g_i is $\{W_i|\delta\}$, the probability that the actual relationship between profit and the complete set of state variables is W_i , which we approximate by g_i.

Note that even if W_i is the actual relationship between profit and the complete set of state variables: $v = W_i(d,\underline{s},\underline{z})$, it is not true that g_i is the actual relationship between profit and the subset of state variables $\underline{s}$: $v \neq g_i(d,\underline{s})$.

To say that there is no "correct" model does not mean that the choice of function g in the abridged modeling strategy is unimportant. The choice of g partially determines the closeness of the Dependence Approximation:

$$\{v|d,\underline{s},\mathcal{S}\} \approx \delta(v - g(d,\underline{s}))$$

and thereby affects the size of the discrepancy between the modeled and authentic profit lotteries. Although no function makes the approximation exact, we want one that makes it reasonably close. An obvious candidate is the conditional expected value:

$$g(d,\underline{s}) = \langle v|d,\underline{s},\mathcal{S}\rangle$$

1.7 The Adequacy of Modeling in Decision Analysis

We perform modeling in decision analysis not for its own sake, but rather to help us with the analysis. The adequacy of the modeling in a decision analysis must therefore be judged according to how well it serves the purposes of the analysis.

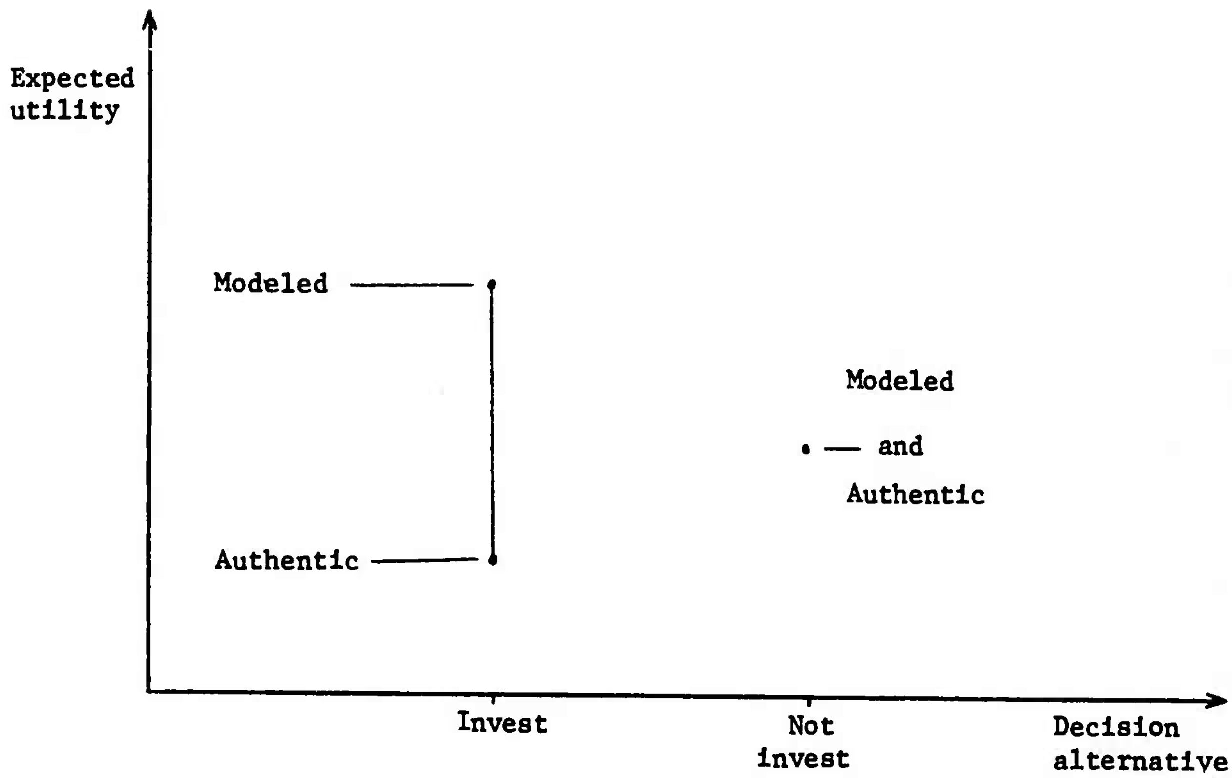
In a decision analysis, we want to identify the optimal decision -- that alternative whose corresponding authentic profit lottery has the highest expected utility. The modeling is adequate if it leads us to the optimal decision.

It is easy to see how modeling might lead us astray. The approximations made in the abridged modeling strategy cause us to misstate the profit lottery and, hence, to misstate the expected utility corresponding to each decision alternative. The danger thus exists that the modeling might cause us to mistakenly identify as optimal an alternative whose authentic expected utility is not highest.

This danger is especially great when much more modeling is required to obtain the profit lottery for one alternative than for another. For example, suppose that the two decision alternatives are to either invest or not invest in a particular project. Obtaining the profit lottery for the first alternative may require extensive modeling while obtaining the profit lottery for the second may require none (zero profit). In this case, the expected utility for the first alternative will be misstated while the expected utility for the second will not, and our decision analysis may then identify the wrong alternative as optimal. (See Figure 1.3).

We can get an indication of the adequacy of the modeling in a decision analysis by using the methodology presented in Chapter 2. The methodology estimates the amount by which we misstate the

Figure 1.3: Effect of Modeling on the Decision



The "not invest" alternative is optimal, but the decision analysis wrongly identifies the "invest" alternative as optimal because the modeling misstates its expected utility.

authentic expected utility for each decision alternative because of the modeling approximations. If the misstatement is shown to be relatively large, we may presume that the modeling is inadequate.

But what can we do if we judge the modeling to be inadequate? The most direct approach is to recalculate the profit lottery for each alternative, trying to get it "closer" to the authentic profit lottery by tightening one or both of the modeling approximations:

1. $\{v|d,\underline{s},\mathcal{S}\} \simeq \delta(v - g(d,\underline{s}))$ (Dependence Approx.)

2. $\{\underline{s}|\mathcal{S}\} \simeq p(\underline{s})$ (Distribution Approx.)

Hoping to minimize the additional computational costs, we might try to improve on these approximations without enlarging the decision tree by choosing a different model g in the first approximation or by choosing different discrete values of the state variables in the second. But these are only remedial measures and we cannot be assured that they will get us closer to the authentic profit lottery. Generally, it is not because we have used the "wrong" model or the "wrong" discrete values of the state variables that the modeling is inadequate; rather, it is because we have been forced by the budgetary constraints into making tenuous modeling approximations.

Hence, to get closer to the authentic profit lottery, we must enlarge the decision tree. We can make the tree "taller" by including additional state variables or "bushier" by making a finer discretization of the aleatory variables or by redesignating some of the non-aleatory variables as aleatory. Or, acknowledging that we are uncertain about the dependence relationship, we can replace the deterministic model g with a multi-point discretization of the conditional distribution $\{v|d,\underline{s},\delta\}$. Any of these measures, of course, increases the computational cost of obtaining the profit lottery.

The methodology presented in Chapter 2 provides us with an indirect way of compensating for the inadequacy of the modeling. The methodology yields an estimated correction term for the expected utility of each alternative, so we can accept as optimal that alternative having the highest corrected expected utility. However, we cannot do so with complete confidence, because the methodology is not exact.

Clearly, the adequacy of the modeling in a decision analysis depends not only on our modeling skill, but also on the computational resources available to us. Although we usually like to think that we can model any situation, we must be willing to concede the possibility that budgetary constraints may prevent us from successfully modeling a particularly complex decision problem. In such a case, the results of the decision analysis may not be meaningful.

CHAPTER 2

A METHODOLOGY: THE EFFECT OF MODELING APPROXIMATIONS

2.1 Introduction

In decision analysis, the profit lottery that we compute by modeling is not identical to the authentic profit lottery, which we ideally want. The discrepancy between the modeled and authentic profit lotteries is due to approximations made in the modeling process. The methodology developed here allows us to quantitatively relate the size of the discrepancy to the modeling approximations made in the analysis.

As an overriding goal, I have tried to keep the methodology simple enough that it can be employed quickly and easily to check the adequacy of the modeling in any decision analysis.

Consider a decision problem defined by decision variable d and outcome variable v. We would like to obtain the authentic profit lottery $\{v|d,\mathcal{S}\}$. To do this, we first specify a set of state variables $\underline{s}$ on which we believe v partially depends. The dependence of v on d and $\underline{s}$ can be expressed as a conditional probability distribution $\{v|d,\underline{s},\mathcal{S}\}$.

If we could assess the authentic distributions $\{\underline{s}|\mathcal{S}\}$ and $\{v|d,\underline{s},\mathcal{S}\}$ and if we could perform the necessary integration, we would obtain the authentic profit lottery via the expansion equation:

$$(2.1) \qquad \{v|d,\delta\} = \int_{\underline{s}} \{v|d,\underline{s},\delta\}\{\underline{s}|\delta\}$$

However, this is not generally possible. Usually, we cannot afford the time and effort to fully assess the distributions nor can we analytically perform the required integration over arbitrary continuous functions.

Consequently, we make two modeling approximations to simplify our task. First, we assume that the probability distribution on $\underline{s}$ can be represented by a discrete probability function $p(\underline{s})$. I call this the Distribution Approximation:

$$(2.2) \qquad \{\underline{s}|\delta\} \simeq p(\underline{s})$$

Included in the Distribution Approximation is the fixation of each non-aleatory variable at a single representative value, in effect approximating its distribution by a single-point discretization.

Secondly, we assume that the dependence of v on d and $\underline{s}$ can be represented by a deterministic function g. I call this the Dependence Approximation:

$$v \simeq g(d,\underline{s})$$

or, equivalently,

$$(2.3) \qquad \{v|d,\underline{s},\delta\} \simeq \delta(v - g(d,\underline{s}))$$

Having made these two modeling approximations, we obtain the familiar block diagram representation of the decision problem and, equivalently, the decision tree, as shown in Figure 2.1. We can then easily compute the modeled profit lottery:

$$(2.4) \qquad \{v|d,\delta\}_m = \sum_{\underline{s}} \delta(v - g(d,\underline{s}))\ p(\underline{s})$$

But, because of the modeling approximations (2.2) and (2.3), the modeled profit lottery $\{v|d,\delta\}_m$ from (2.4) is not identical to the authentic profit lottery $\{v|d,\delta\}$ from (2.1).

In the methodology that follows, we shall quantitatively characterize the modeling approximations and determine their effects on the discrepancy between the profit lotteries. We shall look at the approximations one at a time, starting with the Distribution Approximation.

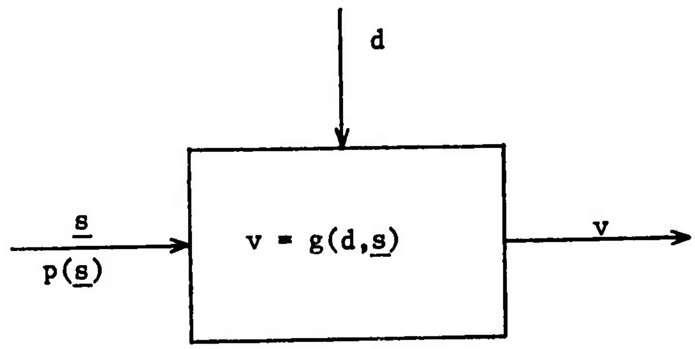
To facilitate the discussion, I shall use subscripts for the outcome variable v to denote the various stages of the development of the methodology (see Figure 2.2):

v_m = the modeled outcome variable

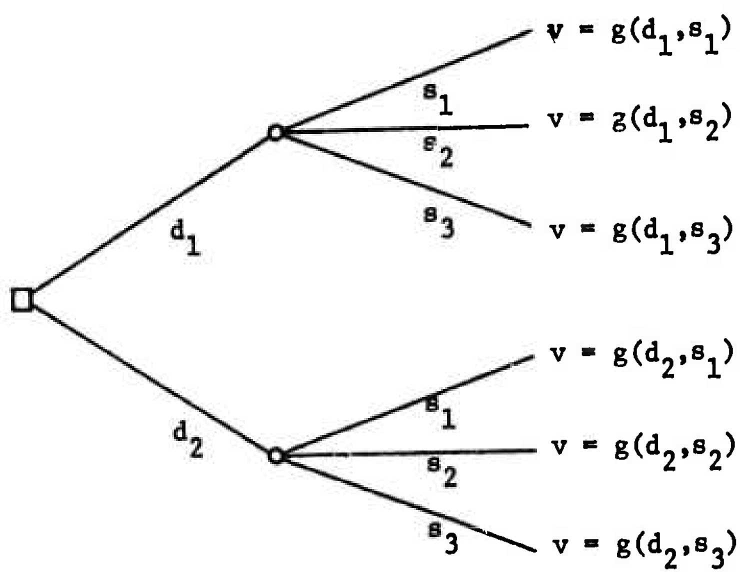
v_x = the outcome variable corrected only for the Distribution Approximation

v_a = the outcome variable corrected for both approximations (i.e., the authentic outcome variable, defined as the variable whose distribution is the authentic profit lottery)

Figure 2.1: Models of the Decision Problem



Block Diagram



Decision Tree

Figures 2.2: Definition of Subscripted Outcome Variables

v_m = Modeled outcome variable

$$\{v_m|d,\&\} = \sum_{\underline{s}} \delta(v - g(d,\underline{s}))\, p(\underline{s})$$

↓

Correction for the Distribution Approximation:

$$\{\underline{s}|\&\} \simeq p(\underline{s})$$

↓

v_x = Outcome variable corrected only for the Distribution Approximation

$$\{v_x|d,\&\} = \int_{\underline{s}} \delta(v - g(d,\underline{s}))\{\underline{s}|\&\}$$

↓

Correction for the Dependence Approximation:

$$\{v|d,\underline{s},\&\} \simeq \delta(v - g(d,\underline{s}))$$

↓

v_a = Authentic outcome variable

$$\{v_a|d,\&\} = \int_{\underline{s}} \{v|d,\underline{s},\&\}\{\underline{s}|\&\}$$

The goal of the methodology is to determine correction terms $\Delta\bar{v}$ and $\Delta\check{v}$ for the mean and variance of the outcome variable:

$$\Delta\bar{v} = \bar{v}_a - \bar{v}_m$$

$$\Delta\check{v} = \check{v}_a - \check{v}_m$$

We shall consider in Sections 2.2, 2.3 and 2.4 the case in which there is only one state variable s and then, in Section 2.5, we shall extend the results to the multivariate case.

2.2 Effect of the Distribution Approximation

Assume that the state variable s is scalar and assume for the moment that $v = g(d,s)$ exactly. We want to determine the effect on the profit lottery of using the discrete distribution $p(s)$ on the state variable rather than the authentic distribution $\{s|\mathcal{S}\}$.

For notational ease, let

n = the state variable whose distribution is $p(s)$

x = the state variable whose distribution is $\{s|\mathcal{S}\}$

The choice of "n" and "x" emphasizes that $p(s)$ is discrete while $\{s|\mathcal{S}\}$ is generally continuous.

We characterize the Distribution Approximation by the differences in the mean and variance of x and n :

$$\Delta\bar{s} = \bar{x} - \bar{n}$$

$$\Delta\check{s} = \check{x} - \check{n}$$

If $\{s|\delta\}$ is known, $\Delta\bar{s}$ and $\Delta\check{s}$ can be computed; otherwise, they can be assessed directly.

Now, noting that $v_m = g(d,n)$ and $v_x = g(d,x)$, we expand both v_m and v_x about $\bar{n}$, using the first three terms of the Taylor Series [11] :

$$(2.5) \qquad v_m \simeq g(d,\bar{n}) + g'(d,\bar{n})(n-\bar{n}) + \tfrac{1}{2}g''(d,\bar{n})(n-\bar{n})^2$$

$$(2.6) \qquad v_x \simeq g(d,\bar{n}) + g'(d,\bar{n})(x-\bar{n}) + \tfrac{1}{2}g''(d,\bar{n})(x-\bar{n})^2$$

where

$$g'(d,\bar{n}) = \frac{\delta}{\delta s} g(d,s)\Big|_{s=\bar{n}}$$

$$g''(d,\bar{n}) = \frac{\delta^2}{\delta s^2} g(d,s)\Big|_{s=\bar{n}}$$

Then, we take the mean and variance of (2.5) and (2.6) (see Appendix A.1) and eliminate third- and fourth-order terms with the following simplifying assumptions (see Appendix A.2) :

$$\text{Third central moment} = \overline{(s-\bar{s})^3} = 0$$

$$\text{Fourth central moment} = \overline{(s-\bar{s})^4} = 3\,\check{s}^2$$

As a result, we have:

$$(2.7) \qquad \bar{v}_m \simeq g(d,\bar{n}) + \tfrac{1}{2}g''(d,\bar{n})\,\check{n}$$

$$(2.8) \qquad \check{v}_m \simeq g'(d,\bar{n})^2\,\check{n} + \tfrac{1}{2}g''(d,\bar{n})^2\,\check{n}^2$$

$$(2.9) \qquad \bar{v}_x \simeq g(d,\bar{n}) + g'(d,\bar{n})\Delta\bar{s} + \tfrac{1}{2}g''(d,\bar{n})(\check{x} + \Delta\bar{s}^2)$$

$$(2.10) \qquad \check{v}_x \simeq g'(d,\bar{n})^2\,\check{x} + \tfrac{1}{2}g''(d,\bar{n})^2\,(\check{x}^2 + 2\check{x}\Delta\bar{s}^2) + 2g'(d,\bar{n})g''(d,\bar{n})\,\check{x}\Delta\bar{s}$$

By subtraction, we get:

$$(2.11) \qquad (\bar{v}_x - \bar{v}_m) \simeq g'(d,\bar{n})\Delta\bar{s} + \tfrac{1}{2}g''(d,\bar{n})(\Delta\check{s} + \Delta\bar{s}^2)$$

$$(2.12) \qquad (\check{v}_x - \check{v}_m) \simeq g'(d,\bar{n})^2\Delta\check{s} + \tfrac{1}{2}g''(d,\bar{n})^2\left[\Delta\check{s}(\check{x}+\check{n}) + 2\check{x}\Delta\bar{s}^2\right] + 2g'(d,\bar{n})g''(d,\bar{n})\,\check{x}\Delta\bar{s}$$

These equations yield the estimated correction terms for the effects of the Distribution Approximation. Equation (2.11) is exact for quadratic g and (2.12) is exact for linear g.

2.3 Effect of the Dependence Approximation

We now consider the effect on the profit lottery of the Dependence Approximation. To compute the profit lottery $\{v|d,\mathcal{S}\}_m$, we assume that v is completely determined by d and s:

$$v = g(d,s)$$

In most real situations, however, we believe that v also depends on other less important, unspecified variables. Hence, by assuming the deterministic relationship, we have suppressed whatever uncertainty we may have about v for fixed values of d and s.

We characterize the Dependence Approximation, then, by a random variable e :

$$v = g(d,s) + e$$

Our residual uncertainty about v given d and s can be expressed as a conditional probability distribution on e , $\{e|d,s,\mathcal{S}\}$. For the purposes of the methodology, we assess the conditional mean $\langle e|d,s,\mathcal{S}\rangle$ and variance $\overset{v}{\langle} e|d,s,\mathcal{S}\rangle$.

Recalling that v_x is corrected for the effects of the Distribution Approximation, we have:

$$v_a = g(d,x) + e = v_x + e \tag{2.13}$$

Taking the mean and variance of this equation (see Appendix A.3), we get:

$$(\bar{v}_a - \bar{v}_x) = \bar{e} \tag{2.14}$$

$$(\check{v}_a - \check{v}_x) = \check{e} + 2\mathrm{cov}(v_x,e) \tag{2.15}$$

These equations yield the correction terms for the effects of the Dependence Approximation.

We can calculate $\bar{e}$, $\check{e}$ and $\text{cov}(v_x,e)$ as follows (see Appendix A.4) :

(2.16) $$\bar{e} = \int_s \langle e|d,s,\delta\rangle\{s|\delta\}$$

(2.17) $$\check{e} = \int_s \left[\left({}^{v}\langle e|d,s,\delta\rangle + \langle e|d,s,\delta\rangle^2\right)\{s|\delta\}\right] - \bar{e}^2$$

(2.18) $$\text{cov}(v_x,e) = \int_s \left[g(d,s)\ \langle e|d,s,\delta\rangle\ \{s|\delta\}\right] - \bar{v}_x\bar{e}$$

Note that each of these calculations requires the integration of a function of d and s (e.g., $\langle e|d,s,\delta\rangle$) over the authentic distribution $\{s|\delta\}$. In most cases, we are unable to perform this integration and must approximate it with a summation over the discrete distribution $p(s)$. This is just another utilization of the Distribution Approximation; that is, for an arbitrary function $f(d,s)$,

$$\int_s f(d,s)\ \{s|\delta\} \simeq \sum_s f(d,s)\ p(s)$$

We can use the results of Section 2.2 to find correction terms for these calculations. Letting:

$$z_x = f(d,x) \quad \text{and} \quad z_n = f(d,n)$$

we see that

$$\bar{z}_x = \int_s f(d,s)\ \{s|\delta\}$$

$$\bar{z}_n = \sum_s f(d,s)\ p(s)$$

Then, we expand z_x and z_n about $\bar{n}$, take the mean and subtract:

$$(2.19) \qquad (\bar{z}_x - \bar{z}_n) \simeq f'(d,\bar{n})\Delta\bar{s} + \tfrac{1}{2}f''(d,\bar{n})(\Delta\check{s} + \Delta\bar{s}^2)$$

This equation, with appropriate substitution for the function $f(d,s)$, yields the estimated correction term for each of the calculations in (2.16), (2.17) and (2.18) to compensate for the use of $p(s)$ rather than $\{s|\mathcal{S}\}$.

Two special cases of the Dependence Approximation merit our attention. As the first special case, suppose that e is independent of s; that is, suppose that our residual uncertainty about v given d and s does not depend on s. Then, we need only assess $\bar{e}$ and $\check{e}$ to characterize the Dependence Approximation:

$$\langle e|d,s,\mathcal{S}\rangle = \langle e|d,\mathcal{S}\rangle = \bar{e}$$

$$\check{\langle} e|d,s,\mathcal{S}\rangle = \check{\langle} e|d,\mathcal{S}\rangle = \check{e}$$

If e is independent of s, it is also independent of v_x, so:

$$\text{cov}(v_x,e) = 0$$

Thus, when e is independent of s, the correction terms for the Dependence Approximation from (2.14) and (2.15) become:

(2.20) $(\bar{v}_a - \bar{v}_x) = \bar{e}$

(2.21) $(\check{v}_a - \check{v}_x) = \check{e}$

As the second special case, suppose that we find it easier to assess our residual uncertainty about the model g in terms of its output v_x rather than its input s. Then, we would assess $\langle e|d,v_x,\delta\rangle$ and $^{\vee}\langle e|d,v_x,\delta\rangle$ instead of $\langle e|d,s,\delta\rangle$ and $^{\vee}\langle e|d,s,\delta\rangle$. In this case, we can always determine $\langle e|d,s,\delta\rangle$ and $^{\vee}\langle e|d,s,\delta\rangle$ by a simple change of argument:

$$\langle e|d,s,\delta\rangle = \langle e|d,v_x,\delta\rangle\Big|_{v_x=g(d,s)}$$

$$^{\vee}\langle e|d,s,\delta\rangle = {}^{\vee}\langle e|d,v_x,\delta\rangle\Big|_{v_x=g(d,s)}$$

However, suppose that our residual uncertainty about the model is proportional to its output v_x in the following sense:

$$\langle e|d,v_x,\delta\rangle = A\, v_x$$

$$^{\vee}\langle e|d,v_x,\delta\rangle = B\, v_x^2$$

where A and B are constants. Such would be the case, for example, if $\{e|d,v_x,\delta\}$ were uniform or triangular over e with a base proportional to v_x. Then, we have (see Appendix A.5) :

$$\bar{e} = \langle e|d,\delta\rangle = A\,\bar{v}_x$$

$$\check{e} = {}^{\vee}\langle e|d,\delta\rangle = (B + A^2)\,\check{v}_x + B\,\bar{v}_x^{\,2}$$

$$\mathrm{cov}(v_x,e) = A\,\check{v}_x$$

So, for this special case, the correction terms for the effects of the Dependence Approximation from (2.14) and (2.15) become:

$$(2.22) \qquad (\bar{v}_a - \bar{v}_x) = A\,\bar{v}_x$$

$$(2.23) \qquad (\check{v}_a - \check{v}_x) = (B + 2A + A^2)\,\check{v}_x + B\bar{v}_x^{\,2}$$

2.4 Combined Effect of Both Modeling Approximations

The correction terms $\Delta\bar{v}$ and $\Delta\check{v}$ for the combined effect of both modeling approximations are simply the sums of the correction terms for each approximation:

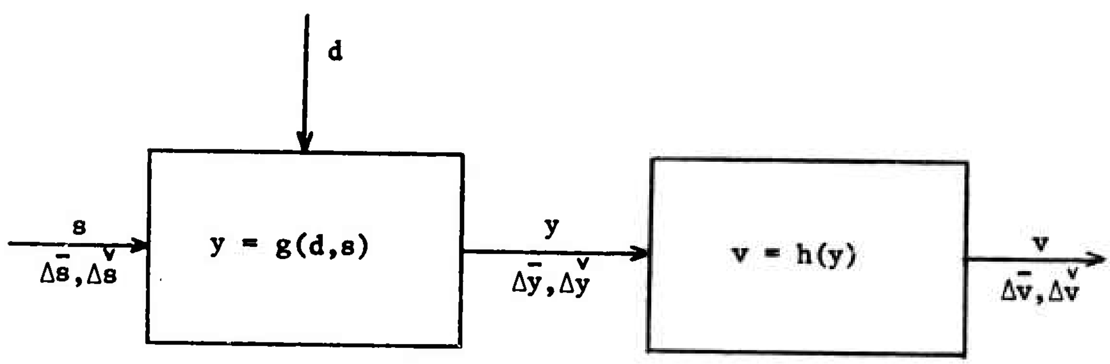
$$(2.24) \qquad \Delta\bar{v} = (\bar{v}_a - \bar{v}_m) = (\bar{v}_a - \bar{v}_x) + (\bar{v}_x - \bar{v}_m)$$

$$(2.25) \qquad \Delta\check{v} = (\check{v}_a - \check{v}_m) = (\check{v}_a - \check{v}_x) + (\check{v}_x - \check{v}_m)$$

Effect of Dependence Approximation (2.14), (2.15)	Effect of Distribution Approximation (2.11), (2.12)

In most cases, the deterministic model that we use to compute the profit lottery is composed of several submodels. We can use the methodology sequentially for each submodel to determine the cumulative effect on the profit lottery of the modeling approximations made in all of the submodels. For example, referring to the model shown in Figure 2.3, we would first use the methodology to determine the correction terms $\Delta\bar{y}$ and $\Delta\check{y}$ for the intermediate variable y. Then, considering y as an input variable to the second submodel, we would use the methodology again

Figure 2.3: Sequential Submodels



to determine the correction terms $\Delta\bar{v}$ and $\Delta\check{v}$ for the outcome variable v.

Note that, if we use the methodology sequentially for several submodels, we need the modeled mean and variance for each intermediate variable. Generally, the computer programs used to calculate the modeled profit lottery can easily be arranged to report these quantities.

Going beyond the profit lottery, we can determine the effect of the modeling approximations on the expected utility and the certain equivalent by considering the utility function $u(v)$ as just another submodel. We want to find the effect on the expected utility $\bar{u}$ of using the modeled profit lottery $\{v|d,\delta\}_m$ rather than the authentic profit lottery $\{v|d,\delta\}$ as the distribution on input variable v. This is the Distribution Approximation once again, which we characterize by the correction terms $\Delta\bar{v}$ and $\Delta\check{v}$. From (2.11), we have:

$$(2.26) \qquad \Delta\bar{u} \simeq u'(\bar{v}_m)\Delta\bar{v} + \tfrac{1}{2}u''(\bar{v}_m)\,(\Delta\check{v} + \Delta\bar{v}^2)$$

Suppose, for example, that the utility function is exponential with risk coefficient γ :

$$u(v) = 1 - e^{-\gamma v}$$

Then:

$$u'(v) = \gamma e^{-\gamma v}$$

$$u''(v) = \gamma^2 e^{-\gamma v}$$

Substituting into (2.26) , we have:

$$\Delta\bar{u} \simeq \gamma e^{-\gamma v_m}(\Delta\bar{v}) - \tfrac{1}{2}\gamma^2 e^{-\gamma v_m} (\Delta\check{v} + \Delta\bar{v}^2)$$

$$\simeq \gamma e^{-\gamma v_m}\left[\Delta\bar{v} - \tfrac{1}{2}\gamma(\Delta\check{v} + \Delta\bar{v}^2)\right]$$

The corrected expected utility is then:

$$\bar{u}_a = \bar{u}_m + \Delta\bar{u}$$

The corrected certain equivalent can be computed from the corrected expected utility.

2.5 Extension to Multivariate Models

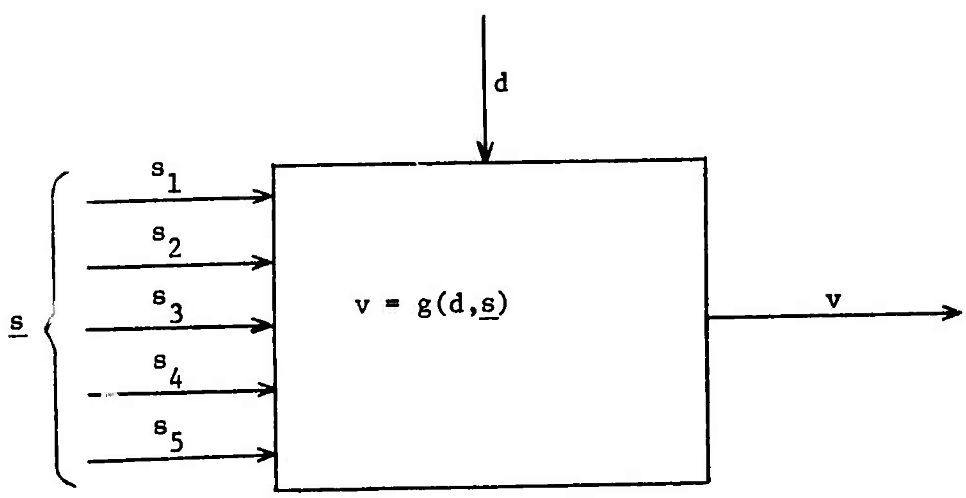
We can extend the methodology developed in the previous sections to decision problems in which there are more than one state variable. However, the methodology is practical only when the state variables are mutually independent.

Consider the multivariate model of a decision problem shown in Figure 2.4. Outcome variable v is dependent on decision variable d and on the state variable vector $\underline{s}$. The dependence is fully expressed by the conditional probability distribution $\{v|d,\underline{s},\mathcal{s}\}_m$, but is approximated for purposes of analysis by a deterministic relationship g :

$$\{v|d,\underline{s},\mathcal{s}\} \simeq \delta(v - g(d,\underline{s}))$$

Also, the authentic distribution on $\underline{s}$, $\{\underline{s}|\mathcal{s}\}$, is approximated by a discrete probability function $p(\underline{s})$.

Figure 2.4: Multivariate Model



Having made these two modeling approximations, we compute the modeled profit lottery:

$$\{v|d,\mathcal{S}\}_m = \sum_{\underline{s}} \delta(v - g(d,\underline{s}))\ p(\underline{s})$$

We want to determine the discrepancy between this modeled profit lottery and the authentic profit lottery:

$$\{v|d,\mathcal{S}\} = \int_{\underline{s}} \{v|d,\underline{s},\mathcal{S}\}\ \{\underline{s}|\mathcal{S}\}$$

We develop the methodology as in the previous sections, looking first at the Distribution Approximation and then at the Dependence Approximation.

Let:

$\underline{n}$ = the state variable vector whose distribution is $p(\underline{s})$

$\underline{x}$ = the state variable vector whose distribution is $\{\underline{s}|\mathcal{S}\}$

We characterize the Distribution Approximation by the differences in the mean and variance of $\underline{x}$ and $\underline{n}$:

$$\Delta\bar{\underline{s}} = \bar{\underline{x}} - \bar{\underline{n}}$$

$$\Delta\check{s} = \check{\underline{x}} - \check{\underline{n}}$$

Then, we expand both $v_m = g(d,\underline{n})$ and $v_x = g(d,\underline{x})$ about $\bar{\underline{n}}$ and take the mean and variance (see Appendix A.6), using simplifying assumptions as before to eliminate third- and fourth-order terms (see

Appendix A.2). As shown in Appendix A.6, if the state variables are not mutually independent, the resulting expressions are not useful because they require more information than we have about the distributions on $\underline{s}$. So, assuming that the state variables are mutually independent, we have:

$$(\bar{v}_x - \bar{v}_m) \simeq \sum_i g_i(d,\bar{\underline{n}})\Delta\bar{s}_i + \tfrac{1}{2}\sum_i g_{ii}(d,\bar{\underline{n}})\Delta\check{s}_i + \tfrac{1}{2}\sum_i\sum_j g_{ij}(d,\bar{\underline{n}})\Delta\bar{s}_i\Delta\bar{s}_j \tag{2.27}$$

$$\begin{aligned}(\check{v}_x - \check{v}_m) \simeq {} & \sum_i g_i(d,\bar{\underline{n}})^2 \Delta\check{s}_i \\ & + \tfrac{1}{2}\sum_i\sum_j g_{ij}(d,\bar{\underline{n}})^2(\check{x}_i\check{x}_j - \check{n}_i\check{n}_j + \check{x}_i\Delta\bar{s}_j{}^2 + \check{x}_j\Delta\bar{s}_i{}^2) \\ & + \sum_i\sum_j\sum_{k\neq j} g_{ij}(d,\bar{\underline{n}})g_{ik}(d,\bar{\underline{n}})\,\check{x}_i\Delta\bar{s}_j\Delta\bar{s}_k \\ & + 2\sum_i\sum_j g_i(d,\bar{\underline{n}})g_{ij}(d,\bar{\underline{n}})\,\check{x}_i\Delta\bar{s}_j\end{aligned} \tag{2.28}$$

where $g_i(d,\bar{\underline{n}}) = \frac{\delta}{\delta s_i} g(d,\underline{s})\Big|_{\underline{s} = \bar{\underline{n}}}$

$$g_{ij}(d,\bar{\underline{n}}) = \frac{\delta^2}{\delta s_i \delta s_j} g(d,\underline{s})\Big|_{\underline{s} = \bar{\underline{n}}}$$

Equations (2.27) and (2.28) yield the estimated correction terms for the effects of the Distribution Approximation.

As a special case, suppose that there are just two independent state variables s_1 and s_2. Then the correction terms from (2.27) and (2.28) reduce to:

$$
\begin{aligned}
(\bar{v}_x - \bar{v}_m) \;\simeq\;& \left.\begin{aligned} & g_1(d,\underline{\bar{n}})\Delta\bar{s}_1 \\ & + \tfrac{1}{2} g_{11}(d,\underline{\bar{n}})(\Delta\check{s}_1 + \Delta\bar{s}_1^{\,2}) \end{aligned}\right\} s_1 \text{ terms} \\
& \left.\begin{aligned} & + g_2(d,\underline{\bar{n}})\Delta\bar{s}_2 \\ & + \tfrac{1}{2} g_{22}(d,\underline{\bar{n}})(\Delta\check{s}_2 + \Delta\bar{s}_2^{\,2}) \end{aligned}\right\} s_2 \text{ terms} \qquad (2.29) \\
& \left. + \tfrac{1}{2} g_{12}(d,\underline{\bar{n}})\Delta\bar{s}_1\Delta\bar{s}_2 \right\} \text{cross term}
\end{aligned}
$$

$$
\begin{aligned}
(\check{v}_x - \check{v}_m) \;\simeq\;& \left.\begin{aligned} & g_1(d,\underline{\bar{n}})^2\Delta\check{s}_1 \\ & + \tfrac{1}{2} g_{11}(d,\underline{\bar{n}})^2\left[\Delta\check{s}_1(\check{x}_1 + \check{n}_1) + 2\check{x}_1\Delta\bar{s}_1^{\,2}\right] \\ & + 2g_1(d,\underline{\bar{n}})g_{11}(d,\underline{\bar{n}})\,\check{x}_1\Delta\bar{s}_1 \end{aligned}\right\} s_1 \text{ terms} \\
& \left.\begin{aligned} & + g_2(d,\underline{\bar{n}})^2\Delta\check{s}_2 \\ & + \tfrac{1}{2} g_{22}(d,\underline{\bar{n}})^2\left[\Delta\check{s}_2(\check{x}_2 + \check{n}_2) + 2\check{x}_2\Delta\bar{s}_2^{\,2}\right] \\ & + 2g_2(d,\underline{\bar{n}})g_{22}(d,\underline{\bar{n}})\,\check{x}_2\Delta\bar{s}_2 \end{aligned}\right\} s_2 \text{ terms}
\end{aligned}
$$

(Expression continued on next page)

$$
(2.30)\qquad \left.\begin{aligned}
&+ g_{12}(d,\bar{\underline{n}})^2(\check{x}_1\check{x}_2 - \check{n}_1\check{n}_2+\check{x}_1\Delta\bar{s}_2{}^2+\check{x}_2\Delta\bar{s}_1{}^2)\\
&+ 2g_{11}(d,\bar{\underline{n}})g_{12}(d,\bar{\underline{n}})\ \check{x}_1\Delta\bar{s}_1\Delta\bar{s}_2\\
&+ 2g_{22}(d,\bar{\underline{n}})g_{12}(d,\bar{\underline{n}})\ \check{x}_2\Delta\bar{s}_1\Delta\bar{s}_2\\
&+ 2g_1(d,\bar{\underline{n}})g_{12}(d,\bar{\underline{n}})\ \check{x}_1\Delta\bar{s}_2\\
&+ 2g_2(d,\bar{\underline{n}})g_{12}(d,\bar{\underline{n}})\ \check{x}_2\Delta\bar{s}_1
\end{aligned}\right\}\ \text{cross terms}
$$

Since the outcome variable is still a scalar, we handle the Dependence Approximation exactly as in Section 2.3. We characterize the approximation by a scalar random variable e :

$$v = g(d,\underline{s}) + e$$

and assess the conditional mean and variance $\langle e|d,\underline{s},\mathcal{s}\rangle$ and $\check{\langle} e|d,\underline{s},\mathcal{s}\rangle$. Then, the correction terms for the effects of the Dependence Approximation from (2.14) and (2.15) are:

$$(2.31)\qquad (\bar{v}_a - \bar{v}_x) = \bar{e}$$

$$(2.32)\qquad (\check{v}_a - \check{v}_x) = \check{e} + 2\mathrm{cov}(v_x,e)$$

where $\bar{e}$, $\check{e}$ and $\mathrm{cov}(v_x,e)$ are calculated as in (2.16) , (2.17) and (2.18).

If we use the discrete distribution $p(\underline{s})$ instead of the authentic distribution $\{\underline{s}|\mathcal{s}\}$ in the calculations of $\bar{e}$, $\check{e}$ and $\mathrm{cov}(v_x,e)$, we can determine correction terms for these calculations as in (2.19). The multivariate counterpart of (2.19) is:

$$(\bar{z}_x - \bar{z}_n) \simeq \sum_i f_i(d,\underline{\bar{n}})\Delta\bar{s}_i + \tfrac{1}{2}\sum_i f_{ii}(d,\underline{\bar{n}})\Delta\check{s}_i + \tfrac{1}{2}\sum_i\sum_j f_{ij}(d,\underline{\bar{n}})\Delta\bar{s}_i\Delta\bar{s}_j \tag{2.33}$$

Finally, the correction terms for the combined effects of both modeling approximations are:

$$\Delta\bar{v} = (\bar{v}_a - \bar{v}_m) = (\bar{v}_a - \bar{v}_x) + (\bar{v}_x - \bar{v}_m)$$

$$\Delta\check{v} = (\check{v}_a - \check{v}_m) = (\check{v}_a - \check{v}_x) + (\check{v}_x - \check{v}_m)$$

Effect of Dependence Approximation (2.31),(2.32)	Effect of Distribution Approximation (2.27),(2.28)

CHAPTER 3

AN EXAMPLE: COMPETITIVE PRICING DECISION

3.1 Introduction

The decision making client plans to produce and sell a new home appliance during the coming year. He knows that he will have one major competitor and that, because the two competing appliances are virtually identical in function, the share of the market each competitor captures will be determined primarily by the relative selling prices. The client is not sure what price his competitor will set.

Furthermore, he is uncertain about the size of the total market. He believes, however, that it will depend strongly on the lower of the two selling prices and on the general economic conditions during the year.

Finally, uncertainty also surrounds the manufacturing costs. The appliance contains an expensive raw material, but the client is uncertain how much the material will cost him and how much of the material is required for each unit.

The client's decision is which selling price he should set for his appliance. Section 3.2 is a description of the decision analysis of this problem and Section 3.3 is an application of the methodology of Chapter 2 to determine how much effect the modeling approximations made in the analysis have on its results.

3.2 Decision Analysis of the Problem

Deterministic phase: In the analysis of this problem, we first construct a deterministic model, as shown in Figure 3.1. The decision variable is the selling price p and the outcome variable is the profit s. Given p and the competitor's selling price y, the client's share of the market z is determined by the Market Split Model:

$$z = \begin{cases} \frac{1}{2}\left(\frac{y}{p}\right)^{k} & p \geq y \\ 1 - \frac{1}{2}\left(\frac{y}{p}\right)^{-k} & p \leq y \end{cases} \qquad \text{where } k = \frac{\ln(10)}{\ln(1.5)} \simeq 5.68$$

See Figure 3.2 for a graph of the Market Split Model.

The Demand Model determines the total market size n given the lower selling price x and an indicator of the general economic conditions f:

$$n = \tfrac{1}{2} f e^{-\left(\frac{x}{225}\right)} \qquad \text{(millions of units)}$$

See Figure 3.3 for a graph of the Demand Model.

The client's total sales q is then:

$$q = z\, n \qquad \text{(millions of units)}$$

His total revenue r is:

$$r = p\, q \qquad \text{(millions of dollars)}$$

Figure 3.1: Deterministic Model

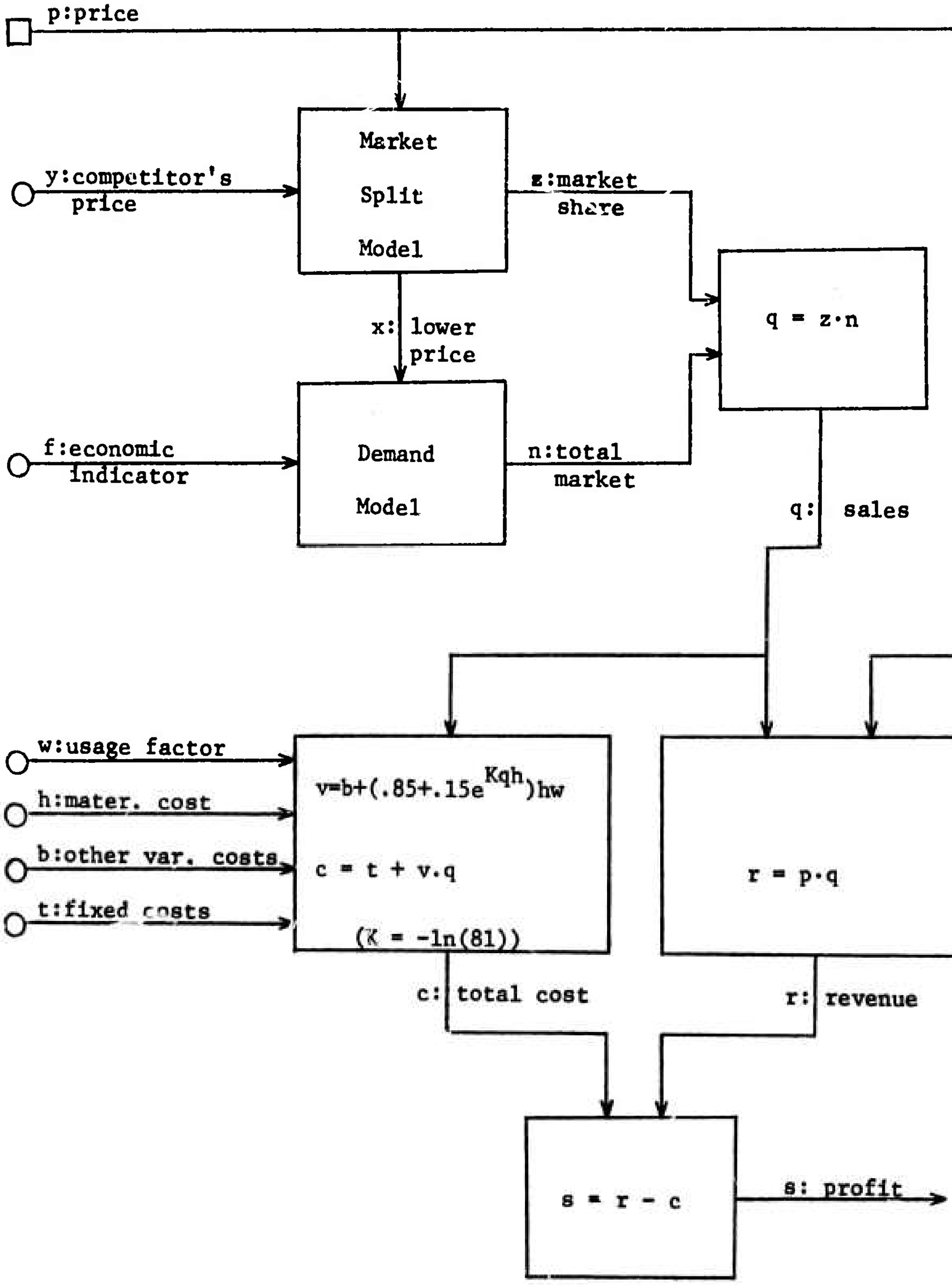


Figure 3.2: Market Split Model

$$z = \begin{cases} 1 - \frac{1}{2}\left(\frac{p}{y}\right)^{k} & \left(\frac{p}{y}\right) \leq 1 \\ \frac{1}{2}\left(\frac{p}{y}\right)^{-k} & \left(\frac{p}{y}\right) \geq 1 \end{cases} \qquad \text{where } k = \frac{\ln(10)}{\ln(1.5)}$$

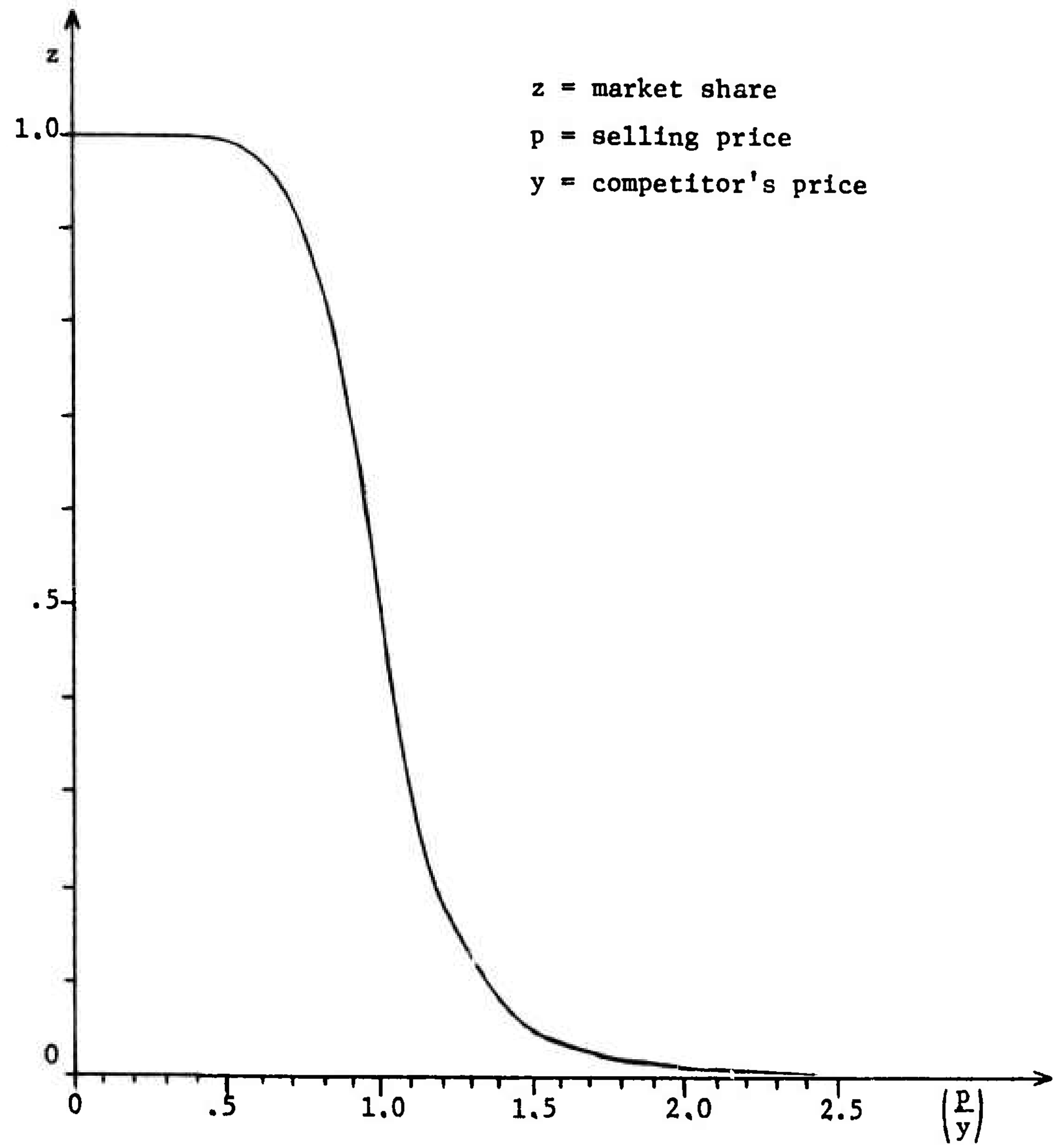
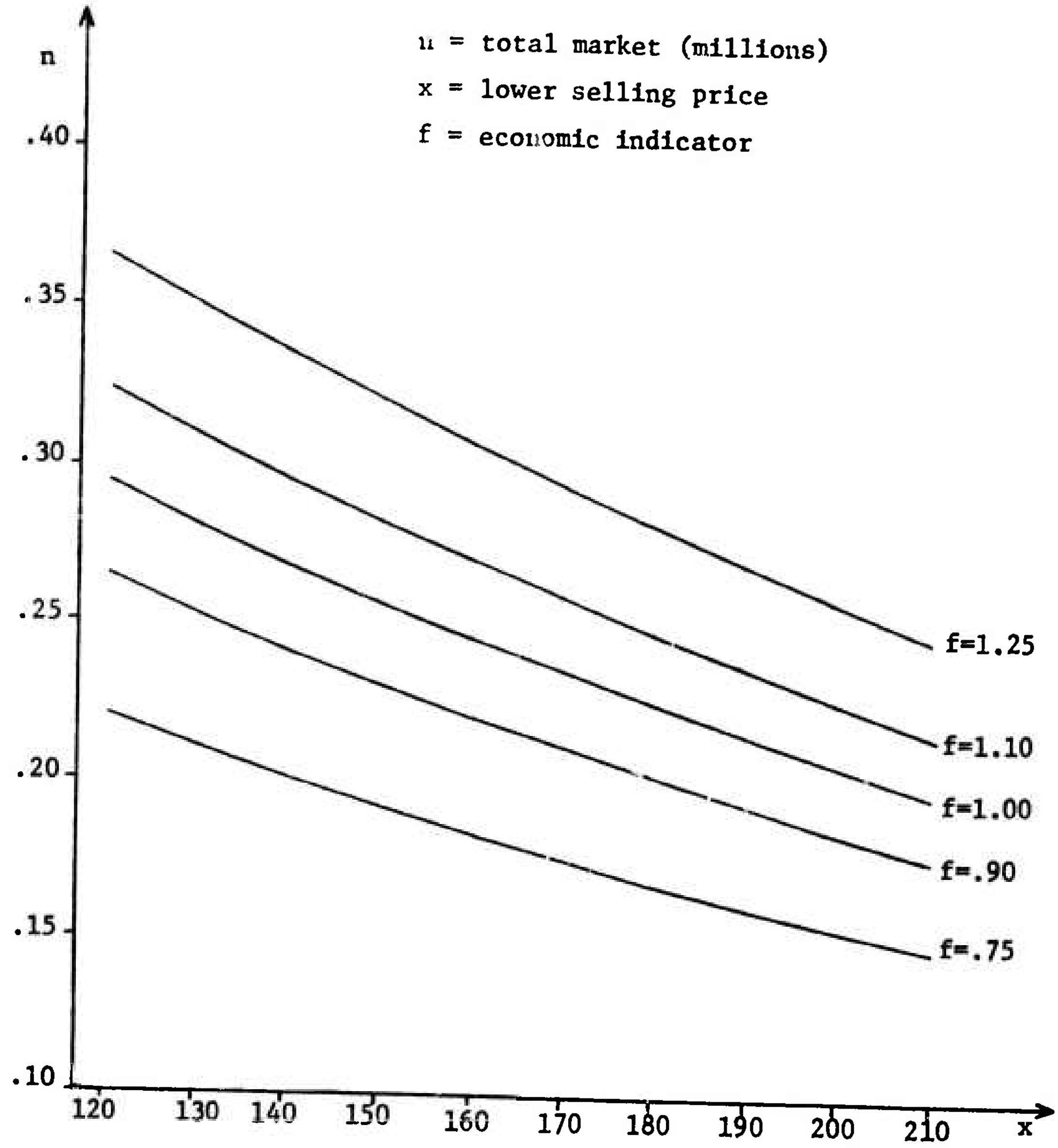


Figure 3.3: Demand Model

$$n = \frac{1}{2} f e^{-\left(\frac{x}{225}\right)}$$



The manufacturing cost per unit v depends on the cost h of the primary raw material, on the per unit usage factor w of that material, on other variable costs b and, because of a volume discount on the primary raw material, on total sales q :

$$v = b + (.85 + .15e^{Kqh})\ h\ w \qquad \text{(dollars)}$$

$$\text{where } K = -\ln(81) \simeq -4.4$$

Then, given fixed costs t , the total cost c is:

$$c = t + v\ q \qquad \text{(millions of dollars)}$$

Finally, profit s is simply:

$$s = r - c \qquad \text{(millions of dollars)}$$

Next, we assign low, nominal and high values for the state variables, as follows:

State variable		Low	Nominal	High
Competitor's price:	y ($)	180	200	250
Economic indicator:	f	0.75	1.00	1.25
Material usage:	w	1.10	1.25	1.40
Raw material cost:	h ($)	30	40	50
Other var. costs:	b ($)	30	40	50
Fixed costs:	t ($$10^6$)	7	9	11

Using the deterministic model and the nominal values of the state variables, we compute the deterministic profit for each of several different values of the decision variable p. The results are shown on Figure 3.4 The highest deterministic profit of s = $6.97 million corresponds to the decision p = 170.

We next calculate the deterministic sensitivity of the profit to each of the state variables by varying each state variable through its range while holding the others at their nominal values. The results of the deterministic sensitivity are shown in Table 3.1. We find that profit is relatively insensitive to the material usage factor w and to the fixed costs t. Profit is highly sensitive to the economic indicator f, although the deterministically optimal decision is not affected by f, as it is by the competitor's price y, the raw material cost h and the other variable costs b.

Probabilistic phase: Because deterministic profit is relatively insensitive to changes in w and t, we designate them as non-aleatory and fix them at their nominal values. We designate the remaining state variables as aleatory variables and, in interviews with the client, assess the probability distribution for each aleatory variable, as shown in Figures 3.5 through 3.8. We then approximate each of these distributions by a discrete distribution as shown in the figures to allow assignment of probabilities to the decision tree. The competitor's price y is discretized into eight steps, the economic indicator f into five steps and the raw material cost h and the other variable costs b into three steps each. The resulting decision tree has 360 "tips" for each

Figure 3.4: Deterministic Results

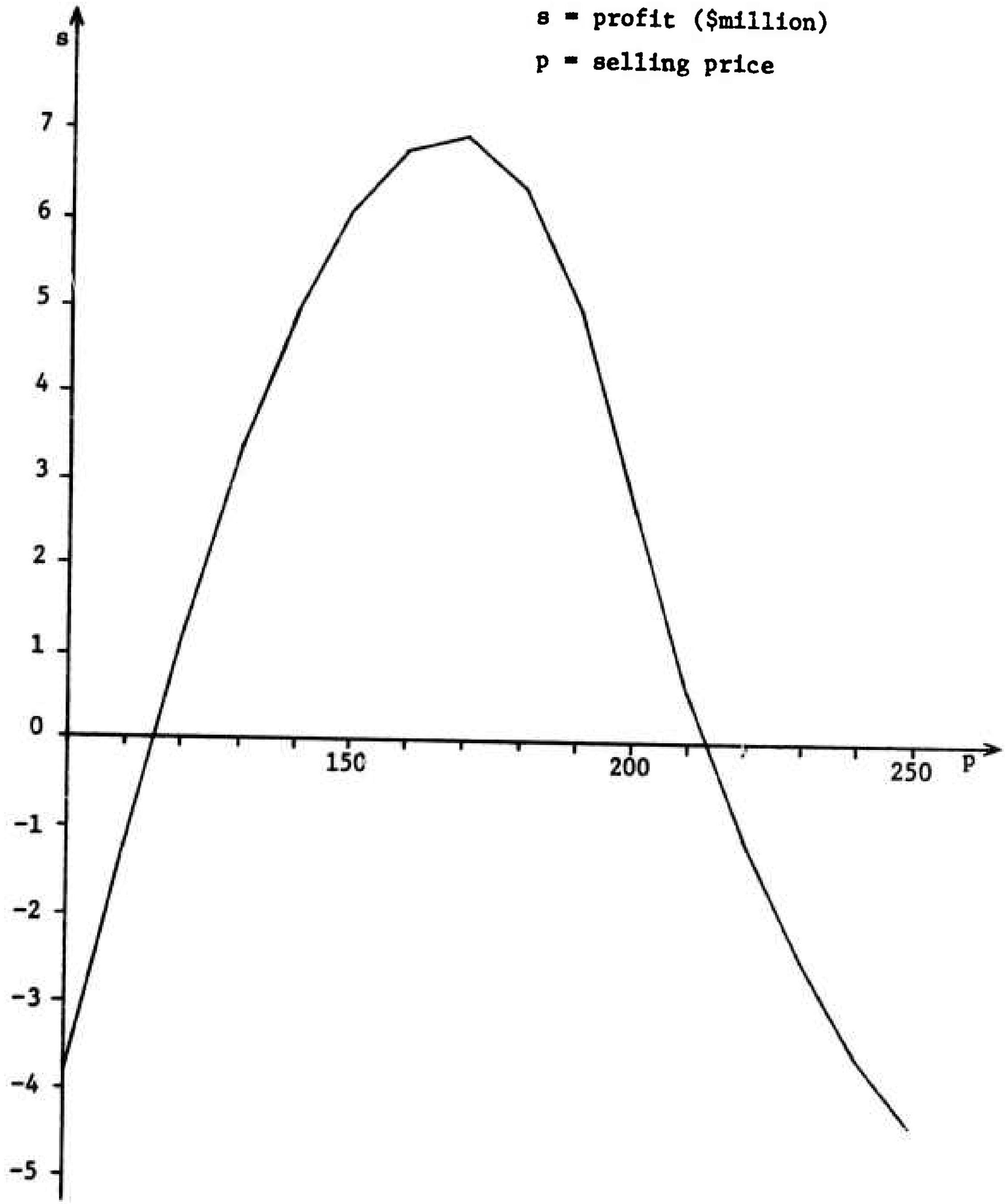


Table 3.1: Deterministic Sensitivity

Sensitivity to y

y	p*	s
180	150	4.75
190	160	5.94
200	170	6.97
210	170	7.96
220	180	8.90
230	190	9.70
240	190	10.53
250	200	11.25

Sensitivity to w

w	p*	s
1.10	170	7.93
1.20	170	7.29
1.25	170	6.97
1.30	170	6.65
1.40	170	6.01

Sensitivity to f

f	p*	s
.75	170	2.86
.90	170	5.32
1.00	170	6.97
1.10	170	8.62
1.25	170	11.10

Sensitivity to b

b	p*	s
30	160	8.96
35	170	7.91
40	170	6.97
45	170	6.03
50	170	5.08

Sensitivity to h

h	p*	s
30	160	9.22
35	160	8.04
40	170	6.97
45	170	5.90
50	170	4.84

Sensitivity to t

t	p*	s
7.0	170	8.97
8.0	170	7.97
9.0	170	6.97
10.0	170	5.97
11.0	170	4.97

Figure 3.5: Probability Distribution on y

y = competitor's price

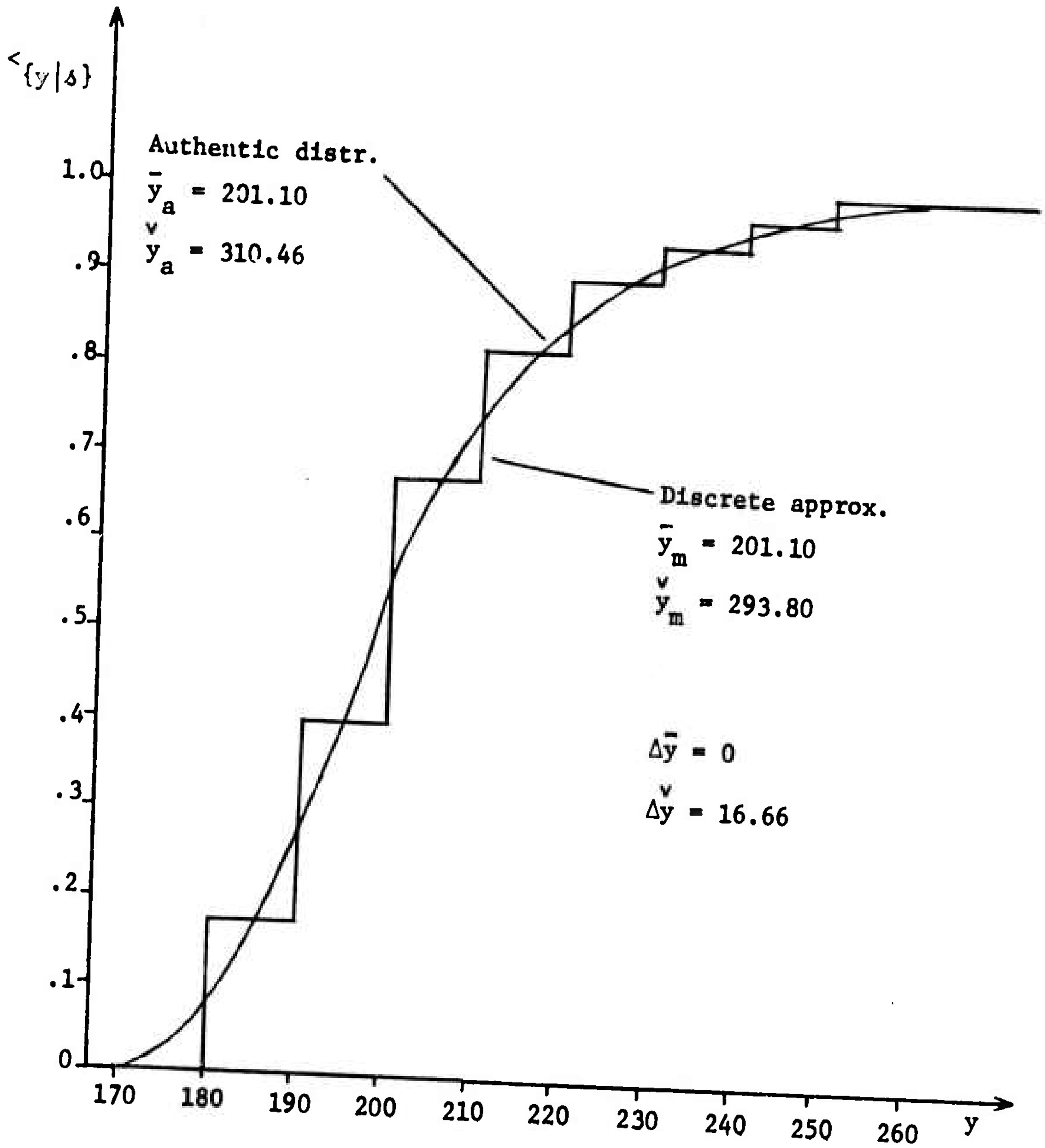


Figure 3.6: Probability Distribution on f

f = economic indicator

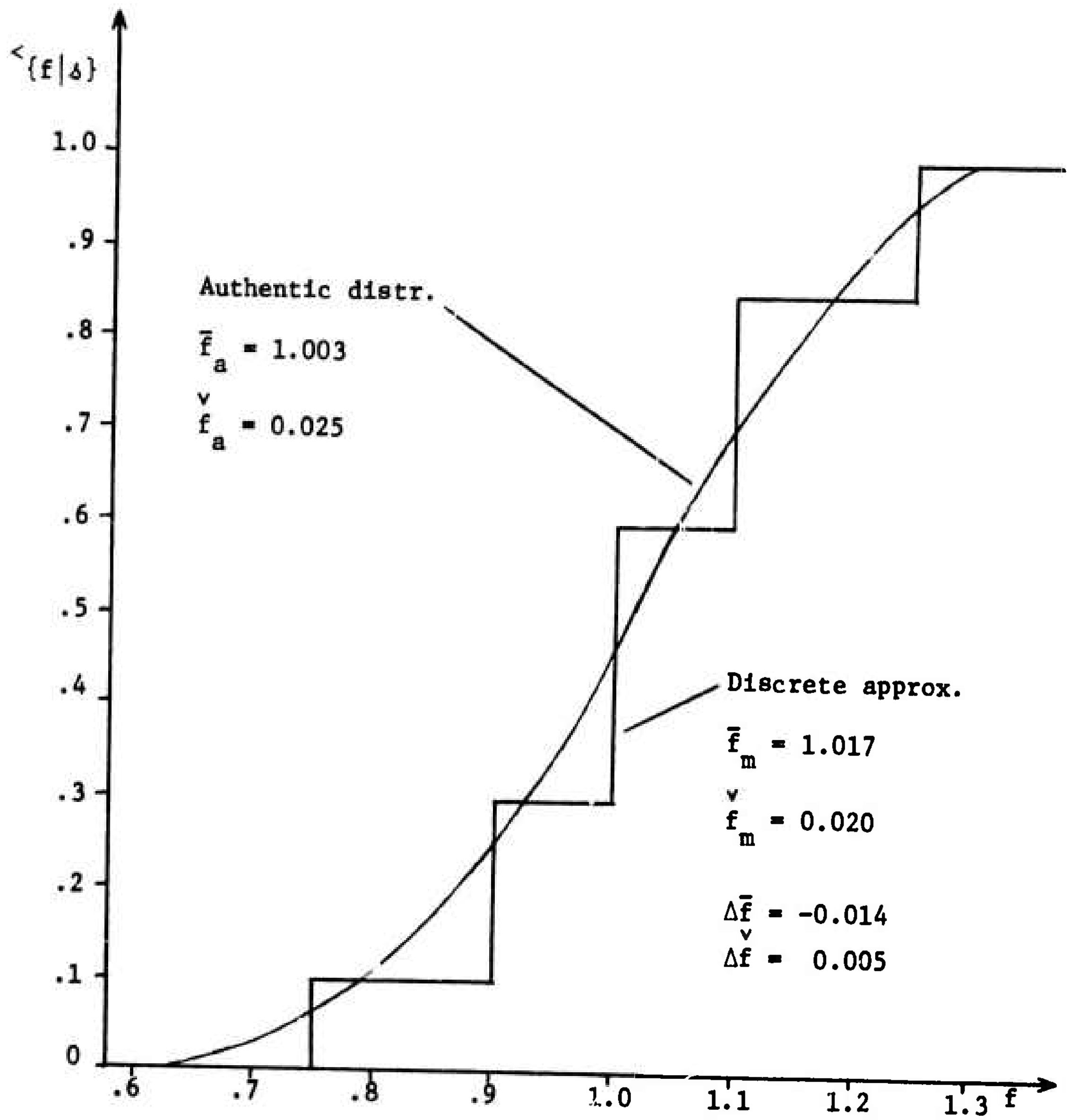


Figure 3.7: Probability Distribution on h

h = raw material cost

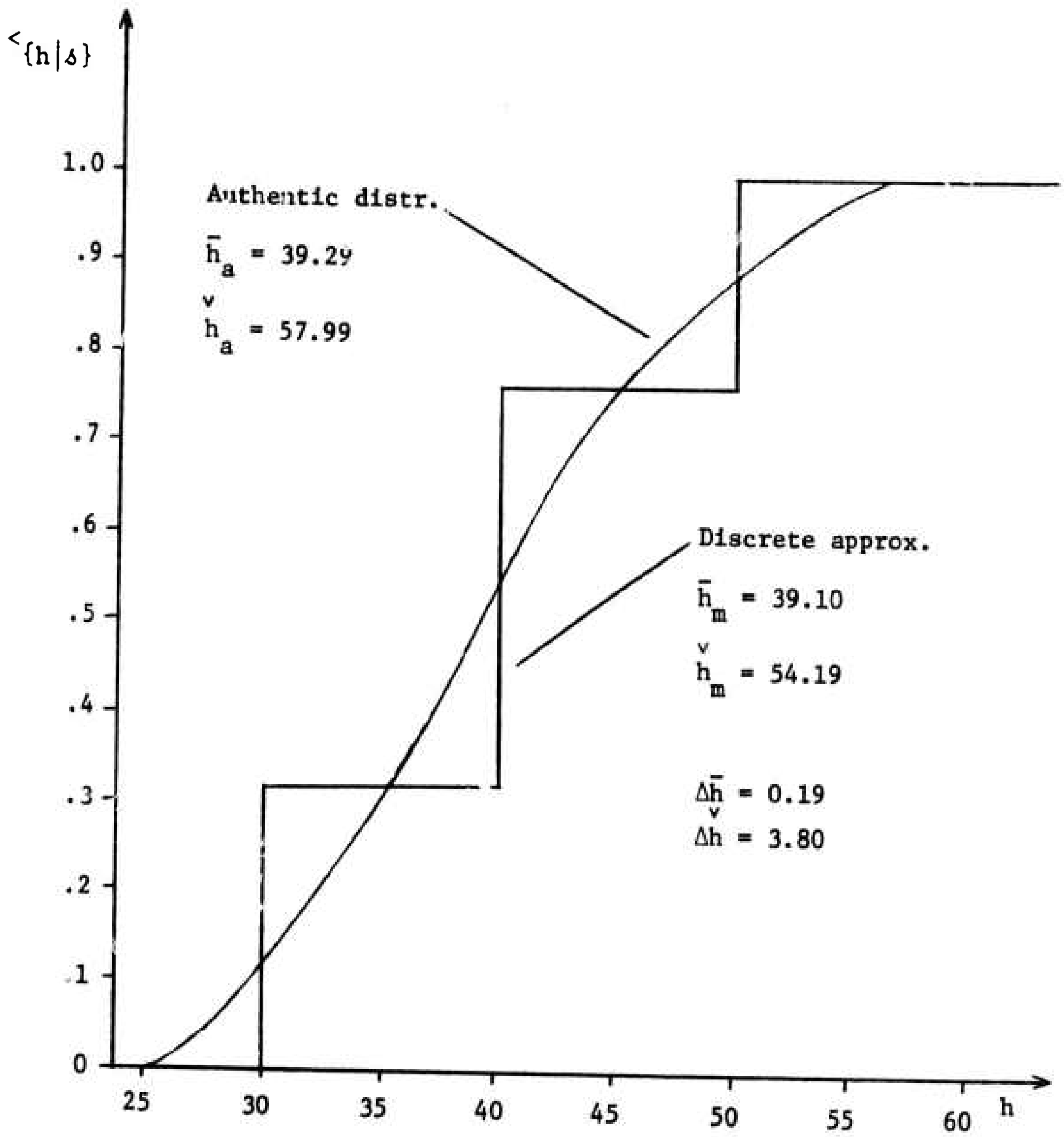
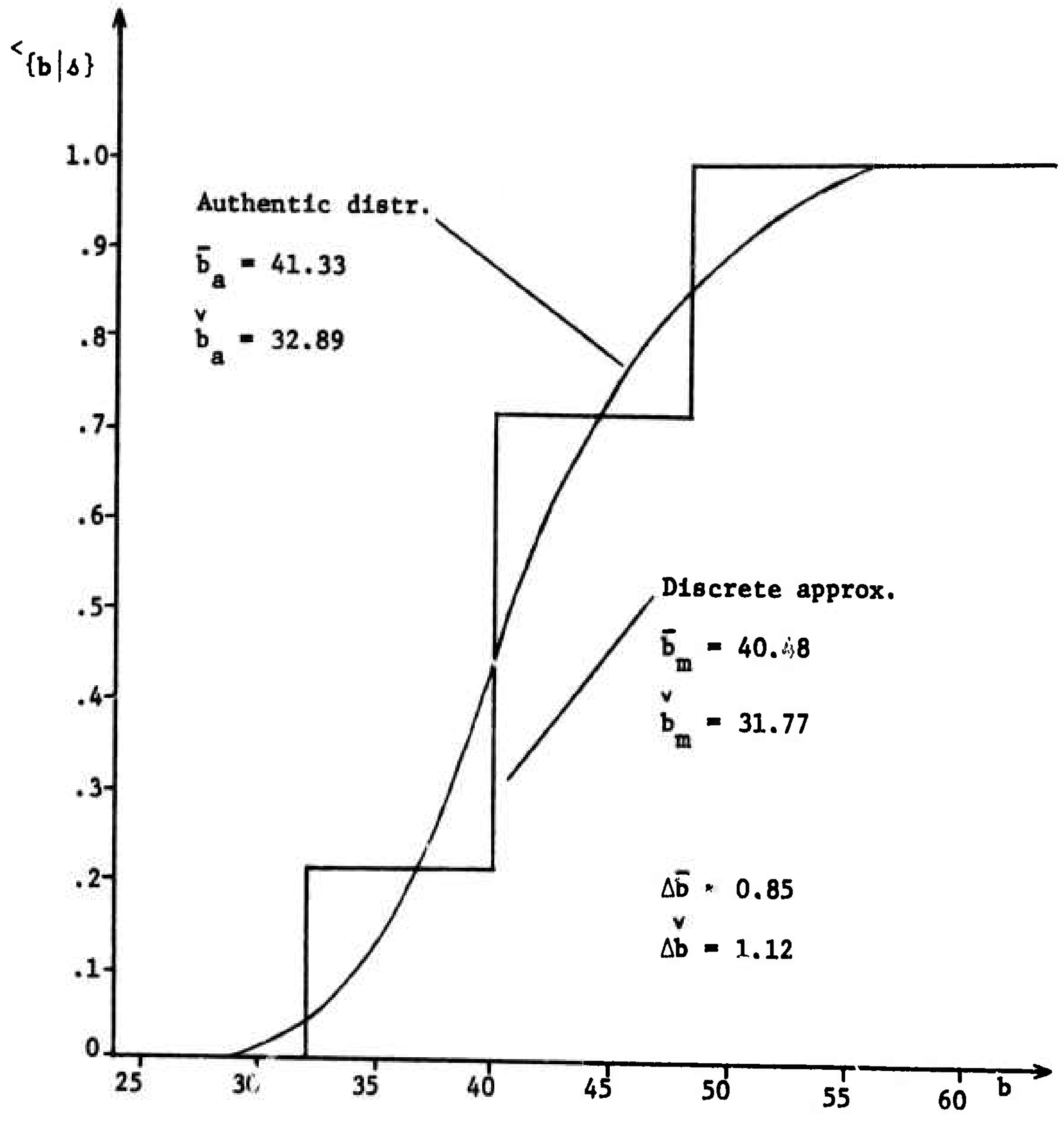


Figure 3.8: Probability Distribution on b

b = other variable costs



decision alternative.

Computations from the decision tree result in the modeled profit lottery $\{s|p,\&\}_m$ for each of several different values of the decision variable p. We determine the relative desirability of each profit lottery by means of the utility function. Interviews with the client reveal that his utility function on profit s is exponential with risk aversion coefficient $.1$:

$$u(s) = 1 - e^{-.1\,s} \quad \text{for } s \text{ in millions of dollars}$$

Using the utility function, we compute the expected utility and the certain equivalent of each profit lottery:

$$\bar{u}_p = \sum_s \{s|p,\&\}_m u(s)$$

$$CE_m(p) = u^{-1}(\bar{u}_p) = -10 \ln(1 - \bar{u}_p)$$

Figure 3.9 is a graph of the certain equivalents. The most preferred profit lottery, shown in Figure 3.10, corresponds to $p = 160$ and has a certain equivalent of $CE_m = \$6.47$ million.

<u>Informational phase:</u> We calculate the probabilistic sensitivity of the certain equivalent to each aleatory variable by holding the variable to each of its values while allowing the other aleatory variables to remain probabilistic. Table 3.2 shows the results of this probabilistic sensitivity analysis. As in the deterministic sensitivity analysis, the optimal decision is not affected by the economic indicator f. The optimal decision is most sensitive to the competitor's price y.

Figure 3.9: Probabilistic Results

CE = certain equivalent ($million)
p = selling price

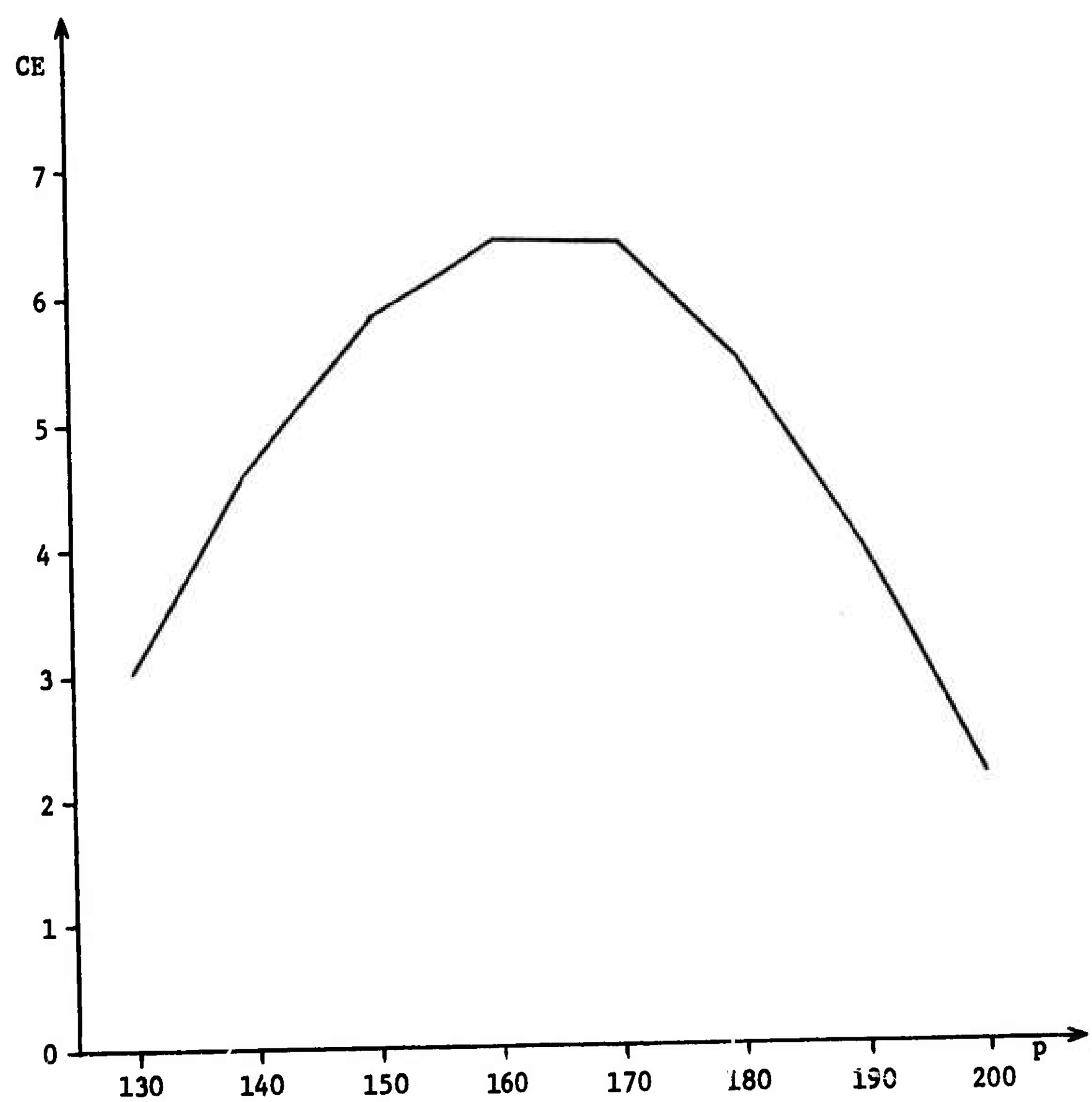


Figure 3.10: Modeled Profit Lottery for p = 160

s = profit ($million)

p = selling price

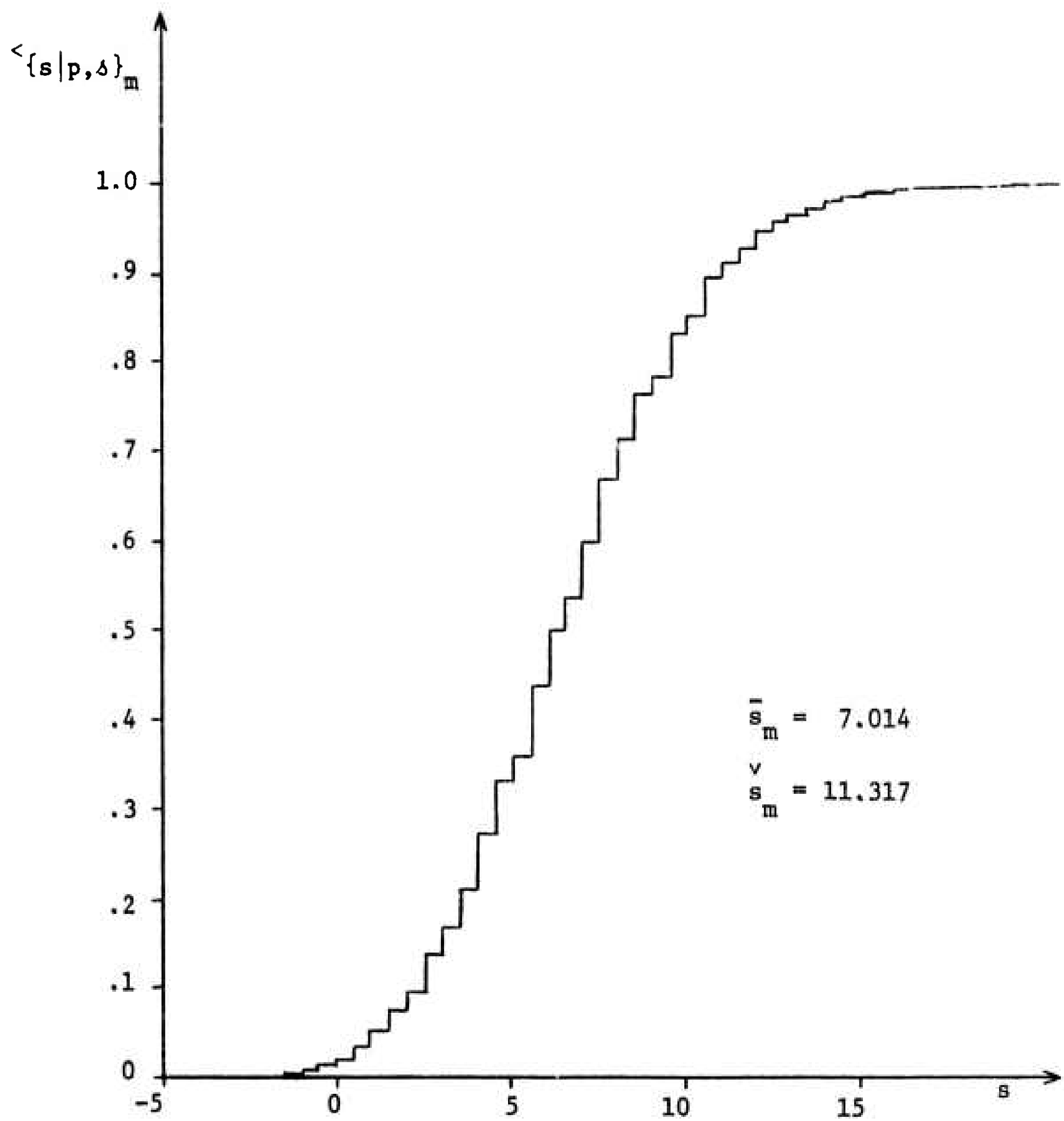


Table 3.2: Probabilistic Sensitivity

Sensitivity to y

y	p*	CE
180	150	4.69
190	160	5.89
200	170	6.91
210	170	7.87
220	180	8.81
230	190	9.61
240	190	10.42
250	200	11.13

Sensitivity to f

f	p*	CE
.75	160	2.51
.90	160	4.87
1.00	160	6.44
1.10	160	8.01
1.25	160	10.37

Sensitivity to h

h	p*	CE
30	160	8.68
40	160	6.41
50	170	4.32

Sensitivity to b

b	p*	CE
32	160	8.25
40	160	6.63
48	170	5.11

Using the results of the probabilistic sensitivity analysis, we compute the value of clairvoyance for each aleatory variable, as follows:

Aleatory variable		Value of clairvoyance
Competitor's price:	y	\$378,000
Economic indicator:	f	0
Raw material cost:	h	\$ 46,000
Other var. costs:	b	\$ 31,000

3.3 Effect of the Modeling Approximations

We can now consider what effect the modeling approximations made in the analysis have on its results. The fundamental question we ask here is how well each of the profit lotteries computed in the analysis (e.g., Figure 3.10) expresses the client's beliefs about the future consequences of his decision or, equivalently, how close the computed profit lottery is to the authentic profit lottery.

The discrepancy between the computed and the autnentic profit lotteries arises from two types of modeling approximations: 1) the approximation of the probability distribution on each state variable by a discrete distribution (Distribution Approximation) and 2) the approximation of the probabilistic relationship between variables by a deterministic relationship (Dependence Approximation).

We can use the methodology developed in Chapter 2 to estimate the discrepancy in the mean and variance of the outcome variable, profit, caused by these approximations.

We characterize the Distribution Approximation for each state variable by the differences in the mean and variance of the variable using the two distributions, as shown in Figures 3.5 through 3.8. These correction terms for all of the state variables are shown in Table 3.3.

We shall trace the effects of the various modeling approximations through the deterministic model shown in Figure 3.1, using the methodology sequentially for each submodel. Figure 3.11 shows the flow of this sequential use.

<u>Effect on market share z of the Distribution Approximation on competitor's price y :</u>

From the Market Split Model, we have for $p \leq y$:

$$z = 1 - \tfrac{1}{2}\left(\frac{y}{p}\right)^{-k} \qquad \text{where} \quad k = \frac{\ln(10)}{\ln(1.5)}$$

Taking derivatives with respect to y :

$$z'(y) = \tfrac{1}{2}\left(\frac{k}{p}\right)\left(\frac{y}{p}\right)^{-(k+1)}$$

$$z''(y) = -\tfrac{1}{2}\,\frac{k(k+1)}{p^2}\left(\frac{y}{p}\right)^{-(k+2)}$$

For $p = 160$, these derivatives, evaluated at $\bar{y}_m = 201.1$ are:

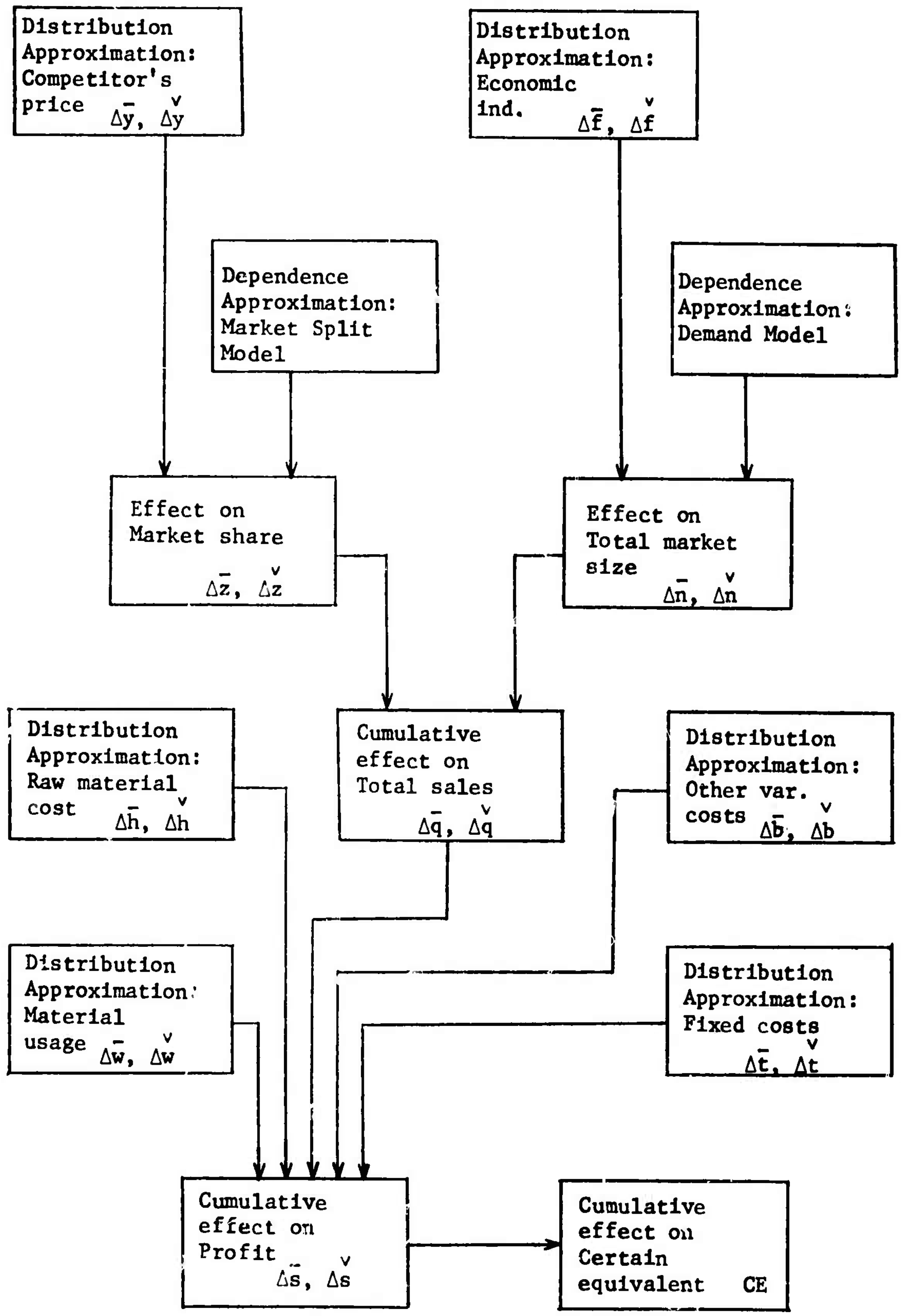
$$z'(\bar{y}_m) = 38.54 \times 10^{-4}$$

$$z''(\bar{y}_m) = -1.28 \times 10^{-4}$$

Table 3.3: Correction Terms for Distribution Approximations

	Authentic	Modeled	Difference
Competitor's price	$\bar{y}_a = 201.10$	$\bar{y}_m = 201.10$	$\Delta\bar{y} = 0$
	$\check{y}_a = 310.46$	$\check{y}_m = 293.80$	$\Delta\check{y} = 16.66$
Economic indicator	$\bar{f}_a = 1.003$	$\bar{f}_m = 1.017$	$\Delta\bar{f} = -0.014$
	$\check{f}_a = 0.025$	$\check{f}_m = 0.020$	$\Delta\check{f} = 0.005$
Raw material cost	$\bar{h}_a = 39.29$	$\bar{h}_m = 39.10$	$\Delta\bar{h} = 0.19$
	$\check{h}_a = 57.99$	$\check{h}_m = 54.19$	$\Delta\check{h} = 3.80$
Other variable costs	$\bar{b}_a = 41.33$	$\bar{b}_m = 40.48$	$\Delta\bar{b} = 0.85$
	$\check{b}_a = 32.89$	$\check{b}_m = 31.77$	$\Delta\check{b} = 1.12$
Material usage factor	$\bar{w}_a = 1.25$	$\bar{w}_m = 1.25$	$\Delta\bar{w} = 0$
	$\check{w}_a = 0.01$	$\check{w}_m = 0$	$\Delta\check{w} = 0.01$
Fixed costs	$\bar{t}_a = 9.0$	$\bar{t}_m = 9.0$	$\Delta\bar{t} =$ [illegible]
	$\check{t}_a = 1.04$	$\check{t}_m = 0$	$\Delta\check{t} = 1.04$

Figure 3.11: Sequential Use of Methodology



Then, noting that $\Delta\bar{y} = 0$, we have from Equations (2.11) and (2.12) the correction terms for the effect on z of the Distribution Approximation:

$$(\bar{z}_x - \bar{z}_m) = \tfrac{1}{2} z''(\bar{y}_m)\, \Delta\check{y}$$

$$(\check{z}_x - \check{z}_m) = z'(\bar{y}_m)^2\, \Delta\check{y} + \tfrac{1}{2} z''(\bar{y}_m)^2\, \Delta\check{y}(\check{y}_a + \check{y}_m)$$

Substituting from Table 3.3 the values of $\check{y}_a$, $\check{y}_m$ and $\Delta\check{y}$:

$$(\bar{z}_x - \bar{z}_m) = (\ 8.33)\ z''(\bar{y}_m)$$

$$(\check{z}_x - \check{z}_m) = (16.66)\ z'(\bar{y}_m)^2 + (5033.5)\ z''(\bar{y}_m)^2$$

For $p = 160$, these correction terms are:

$$(\bar{z}_x - \bar{z}_m) = -.0011$$

$$(\check{z}_x - \check{z}_m) = 3.30 \times 10^{-4}$$

Effect on market share z of the Dependence Approximation in the Market Split Model:

The model asserts that the split of the market is determined solely by the two selling prices. The client believes, however, that the model does not fully represent his uncertainty about the market split. For instance, he believes that other factors, such as advertising and brand loyalty may also influence the market split, although to a lesser extent than the relative prices.

We characterize this Dependence Approximation by the random variable e_z :

$$z_a = z_x + e_z$$

The client believes that the model is unbiased but that his residual uncertainty has a standard deviation of about 8% of the value of z given by the model for $z \leq \frac{1}{2}$ and a symmetric amount for $z \geq \frac{1}{2}$. Therefore, he assesses:

$$\langle e_z | p, z_x, \mathcal{S} \rangle = 0$$

$$^{v}\langle e_z | p, z_x, \mathcal{S} \rangle = \begin{cases} \frac{1}{150} z_x^2 & 0 \leq z_x \leq \frac{1}{2} \\ \frac{1}{150}(1 - z_x)^2 & \frac{1}{2} \leq z_x \leq 1 \end{cases}$$

Since the conditional mean is zero, $\bar{e}_z = 0$ and $\text{cov}(z_x, e_z) = 0$. Then, calculating $\check{e}_z$ for $z \geq \frac{1}{2}$:

$$\check{e}_z = \int_z {}^{v}\langle e_z | p, z_x, \mathcal{S} \rangle \{z_x | p, \mathcal{S}\}$$

$$= \int_z \frac{1}{150}(1 - z_x)^2 \{z_x | p, \mathcal{S}\}$$

$$= \frac{1}{150}(1 - 2\bar{z}_x + \overline{z_x^2})$$

$$= \frac{1}{150}(1 - 2\bar{z}_x + \check{z}_x + \bar{z}_x^2)$$

From (2.14) and (2.15) , the correction terms for the effect on z of the Dependence Approximation are:

$$(\bar{z}_a - \bar{z}_x) = \bar{e}_z = 0$$

$$(\check{z}_a - \check{z}_x) = \check{e}_z + 2\text{cov}(z_x, e_z) = \check{e}_z$$

For $p = 160$, the modeled mean and variance of z are:

$$\bar{z}_m = .847$$

$$\check{z}_m = 38.45 \times 10^{-4}$$

So, for $p = 160$:

$$\bar{z}_x = \bar{z}_m + (\bar{z}_x - \bar{z}_m) = .846$$

$$\check{z}_x = \check{z}_m + (\check{z}_x - \check{z}_m) = 41.75 \times 10^{-4}$$

Then, the correction terms for the effect on z of the Dependence Approximation are for $p = 160$:

$$(\bar{z}_a - \bar{z}_x) = 0$$

$$(\check{z}_a - \check{z}_x) = \frac{1}{150} (1 - 2\bar{z}_x + \check{z}_x + \bar{z}_x^{\,2}) = 1.86 \times 10^{-4}$$

Combined effect on market share z of both approximations

The correction terms for the combined effect on z of both modeling approximations are:

$$\Delta \bar{z} = (\bar{z}_a - \bar{z}_x) + (\bar{z}_x - \bar{z}_m)$$

$$\Delta \check{z} = (\check{z}_a - \check{z}_x) + (\check{z}_x - \check{z}_m)$$

For $p = 160$, these correction terms are:

$$\Delta \bar{z} = -.0011 + 0 = -.0011$$

$$\Delta \check{z} = 3.30 \times 10^{-4} + 1.86 \times 10^{-4} = 5.16 \times 10^{-4}$$

Then, the estimated authentic mean and variance of z for $p = 160$ are:

$$\bar{z}_a = \bar{z}_m + \Delta\bar{z} = .846$$

$$\check{z}_a = \check{z}_m + \Delta\check{z} = 43.61 \times 10^{-4}$$

We see that, for $p = 160$, the analysis overstates the mean of z by about 0.1% and understates the variance of z by about 12% and the standard deviation by about 6%.

<u>Effect on market size n of the Distribution Approximation on the economic indicator f :</u>

From the Demand Model, we have:

$$n = \tfrac{1}{2} f \, e^{-\left(\frac{x}{225}\right)}$$

Taking derivatives with respect to f :

$$n'(f) = \tfrac{1}{2} e^{-\left(\frac{x}{225}\right)}$$

$$n''(f) = 0$$

For $p = 160$, $x = 160$ so:

$$n'(f) = .246 \qquad \text{for all } f$$

Noting that $n''(f) = 0$, we have the correction terms for the effect on n of the Distribution Approximation from (2.11) and (2.12) :

$$(\bar{n}_x - \bar{n}_m) = n'(\bar{f}_m)\,\Delta\bar{f}$$

$$(\check{n}_x - \check{n}_m) = n'(\bar{f}_m)^2\,\Delta\check{f}$$

Substituting from Table 3.3 the values of $\Delta\bar{f}$ and $\Delta\check{f}$:

$$(\bar{n}_x - \bar{n}_m) = (-0.014)\; n'(\bar{f}_m)$$

$$(\check{n}_x - \check{n}_m) = (0.005)\; n'(\bar{f}_m)^2$$

For $p = 160$, these correction terms are:

$$(\bar{n}_x - \bar{n}_m) = -.0035$$

$$(\check{n}_x - \check{n}_m) = 3.24 \times 10^{-4}$$

Effect on market size n of the Dependence Approximation in the Demand Model:

The Demand Model claims that only the lower selling price and the general economic conditions affect the size of the total market. As with the Market Split Model, the client believes that other factors may also influence the total market size and that, therefore, the model understates his uncertainty about the market.

We characterize this Dependence Approximation by the random variable e_n :

$$n_a = n_x + e_n$$

The client assesses his residual uncertainty as:

$$\langle e_n | x, f, \mathcal{S} \rangle = 0$$

$$^v\langle e_n | x, f, \mathcal{S} \rangle = (1.5 \times 10^{-5})(x - 120)\ f^2$$

Again, since the conditional mean is zero, $\bar{e}_n = 0$ and $\text{cov}(n_x, e_n) = 0$. Then, calculating $\check{e}_n$:

$$\check{e}_n = \int_f {}^v\langle e_n | x, f, \mathcal{S} \rangle \{f | \mathcal{S}\}$$

$$= (1.5 \times 10^{-5})\ (x - 120) \int_f f^2 \{f | \mathcal{S}\}$$

$$= (1.5 \times 10^{-5})\ (x - 120)\ \overline{f_a^{\,2}}$$

$$= (1.5 \times 10^{-5})\ (x - 120)(\check{f}_a + \bar{f}_a^{\,2})$$

Substituting from Table 3.3 the values of $\check{f}_a$ and $\bar{f}_a$:

$$\check{e}_n = (1.55 \times 10^{-5})\ (x - 120)$$

The correction terms for the effect on n of the Dependence Approximation from (2.14) and (2.15) are:

$$(\bar{n}_a - \bar{n}_x) = \bar{e}_n = 0$$

$$(\check{n}_a - \check{n}_x) = \check{e}_n + 2\text{cov}(n_x, e_n) = \check{e}_n$$

For $p = 160$, these correction terms are:

$$(\bar{n}_a - \bar{n}_x) = 0$$

$$(\check{n}_a - \check{n}_x) = 6.19 \times 10^{-4}$$

Combined effect on market size n of both approximations:

The correction terms for the combined effect on n of both modeling approximations are:

$$\Delta\bar{n} = (\bar{n}_a - \bar{n}_x) + (\bar{n}_x - \bar{n}_m)$$

$$\Delta\check{n} = (\check{n}_a - \check{n}_x) + (\check{n}_x - \check{n}_m)$$

For $p = 160$, these correction terms are:

$$\Delta\bar{n} = -.0035 \quad + 0 \quad = -.0035$$

$$\Delta\check{n} = 3.24 \times 10^{-4} + 6.19 \times 10^{-4} = 9.43 \times 10^{-4}$$

The modeled mean and variance of n for $p = 160$ are:

$$\bar{n}_m = .250$$

$$\check{n}_m = 11.95 \times 10^{-4}$$

So, the estimated authentic mean and variance for $p = 160$ are:

$$\bar{n}_a = \bar{n}_m + \Delta\bar{n} = .246$$

$$\check{n}_a = \check{n}_m + \Delta\check{n} = 21.38 \times 10^{-4}$$

We see that, for $p = 160$, the analysis overstates the mean of n by about 1.4% and understates the variance of n by about 44% and the standard deviation by about 25%.

Cumulative effect of the modeling approximations on total sales q :

Now, considering the output variables of the Market Split and Demand Models, market share z and total market size n, as input variables to the next submodel, we can determine the cumulative effect on total sales q of the various modeling approximations. These approximations are characterized by the correction terms $\Delta\bar{z}$, $\Delta\check{z}$, $\Delta\bar{n}$ and $\Delta\check{n}$ that we have just computed.

From the model, we have:

$$q = z\,n$$

This is an exact relationship, so there is no Dependence Approximation for this submodel; we need worry about only the Distribution Approximation on input variables z and n. Taking partial derivatives:

$$q_z(z,n) = n \qquad q_{zz}(z,n) = 0$$

$$q_n(z,n) = z \qquad q_{nn}(z,n) = 0$$

$$q_{zn}(z,n) = 1$$

For $p = 160$:

$$q_z(\bar{z}_m,\bar{n}_m) = .250$$

$$q_n(\bar{z}_m,\bar{n}_m) = .847$$

Noting that q_{zz} and q_{nn} are zero, we have the correction terms for the effect on q of the Distribution Approximation from (2.29) and (2.30) :

$$(\bar{q}_x - \bar{q}_m) = q_z(\bar{z}_m,\bar{n}_m)\Delta\bar{z} + q_n(\bar{z}_m,\bar{n}_m)\Delta\bar{n}$$

$$+ \tfrac{1}{2}q_{zn}(\bar{z}_m,\bar{n}_m)\Delta\bar{z}\Delta\bar{n}$$

$$(\check{q}_x - \check{q}_m) = q_z(\bar{z}_m,\bar{n}_m)^2\Delta\check{z} + q_n(\bar{z}_m,\bar{n}_m)^2\Delta\check{n}$$

$$+ q_{zn}(\bar{z}_m,\bar{n}_m)^2(\check{z}_a\check{n}_a - \check{z}_m\check{n}_m + \check{z}_a\Delta\bar{n}^2 + \check{n}_a\Delta\bar{z}^2)$$

$$+ 2q_z(\bar{z}_m,\bar{n}_m)q_{zn}(\bar{z}_m,\bar{n}_m)\ \check{z}_a\Delta\bar{n}$$

$$+ 2q_n(\bar{z}_m,\bar{n}_m)q_{zn}(\bar{z}_m,\bar{n}_m)\ \check{n}_a\Delta\bar{z}$$

Since there is no Dependence Approximation:

$$\Delta\bar{q} = (\bar{q}_a - \bar{q}_x) + (\bar{q}_x - \bar{q}_n) = (\bar{q}_x - \bar{q}_m)$$

$$\Delta\check{q} = (\check{q}_a - \check{q}_x) + (\check{q}_x - \check{q}_m) = (\check{q}_x - \check{q}_m)$$

For $p = 160$, these correction terms are:

$$\Delta\bar{q} = -.0032$$

$$\Delta\check{q} = 7.02 \times 10^{-4}$$

The modeled mean and variance of q for $p = 160$ are:

$$\bar{q}_m = .212$$

$$\check{q}_m = 11.01 \times 10^{-4}$$

So, the estimated authentic mean and variance of total sales q for $p = 160$ are:

$$\bar{q}_a = \bar{q}_m + \Delta\bar{q} = .209$$

$$\check{q}_a = \check{q}_m + \Delta\check{q} = 18.03 \times 10^{-4}$$

We see that, for $p = 160$, the analysis overstates the mean of q by about 1.5% and understates the variance of q by about 39% and the standard deviation by about 22%.

Cumulative effect of modeling approximations on profit s :

In the model, profit s is given as the following function of selling price p, total sales q, fixed costs t, other variable costs b, raw material cost h and material usage factor w :

$$s = p\,q - t - q\,(b + .85hw + .15hw\, e^{Kqw})$$

$$= q\,(p - b - .85hw - .15hw\, e^{Kqw}) - t$$

where $K = -\ln(81)$

This is an exact relationship, so there is no Dependence Approximation. Taking partial derivatives:

$$s_b = -q$$

$$s_h = -wq(.85 + .15\, e^{Kqw})$$

$$s_w = -hw(.85 + .15(1 + Kq)\, e^{Kqw})$$

$$s_q = p - b - hw(.85 + .15(1 + Kqw)\, e^{Kqw})$$

$$s_t = -1$$

$$s_{bb} = 0$$

$$s_{hh} = 0$$

$$s_{ww} = -.15\, Khq^2\,(1 + Kq)\, e^{Kqw}$$

$$s_{qq} = -.15\, Khw^2\,(2 + Kqw)\, e^{Kqw}$$

$$s_{tt} = 0$$

$$s_{bh} = 0$$

$$s_{bw} = 0$$

$$s_{bq} = -1$$

$s_{bt} = 0$

$s_{hw} = -q(.85 + .15(1 + Kqw)\ e^{Kqw})$

$s_{hq} = -w(.85 + .15(1 + Kqw)\ e^{Kqw})$

$s_{ht} = 0$

$s_{wq} = -h(.85 + .15(1 + 2Kq + Kqw + K^2q^2w)\ e^{Kqw})$

$s_{wt} = 0$

$s_{qt} = 0$

For notational compactness, let:

$$s(.) = s(p, \bar{q}_m, \bar{w}_m, \bar{h}_m, \bar{b}_m, \bar{t}_m)$$

Then, for $p = 160$, the partial derivatives evaluated at the modeled mean are:

$s_b(.) = -.212$

$s_h(.) = -.238$

$s_w(.) = -7.072$

$s_q(.) = 78.353$

$s_{ww}(.) = .025$

$s_{qq}(.) = 10.500$

$s_{hw}(.) = -.178$

$s_{hq}(.) = -1.053$

$s_{wq}(.) = -31.509$

Noting that $\Delta\bar{y}$, $\Delta\bar{w}$ and $\Delta\bar{t}$ are all zero, we have the correction terms for the effect on s of the Distribution Approximation from (2.27) and (2.28) :

$$\begin{aligned}(\bar{s}_x - \bar{s}_m) = {} & s_b(.)\Delta\bar{b} + s_h(.)\Delta\bar{h} + s_q(.)\Delta\bar{q} \\ & + \tfrac{1}{2}\, s_{ww}(.)\Delta\check{w} + \tfrac{1}{2}\, s_{qq}(.)(\Delta\check{q} + \Delta\bar{q}^{\,2}) \\ & + s_{bq}(.)\Delta\bar{b}\Delta\bar{q} + s_{hq}(.)\Delta\bar{h}\Delta\bar{q}\end{aligned}$$

$$\begin{aligned}(\check{s}_x - \check{s}_m) = {} & s_b(.)^2\Delta\check{b} + s_h(.)^2\Delta\check{h} + s_w(.)^2\Delta\check{w} \\ & + s_q(.)^2\Delta\check{q} + s_t(.)^2\Delta\check{t} + \tfrac{1}{2}s_{ww}(.)^2\Delta\check{w}^2 \\ & + \tfrac{1}{2}\, s_{qq}(.)^2(\check{q}_a^{\,2} - \check{q}_m^{\,2} + 2\check{q}_a\Delta\bar{q}^2) \\ & + s_{bq}(.)^2(\check{b}_a\check{q}_a - \check{b}_m\check{q}_m + \check{b}_a\Delta\bar{q}^2 + \check{q}_a\Delta\bar{b}^2) \\ & + s_{hq}(.)^2(\check{h}_a\check{q}_a - \check{h}_m\check{q}_m + \check{h}_a\Delta\bar{q}^2 + \check{q}_a\Delta\bar{h}^2) \\ & + s_{hw}(.)^2(\check{h}_a + \Delta\bar{h}^2)\Delta\check{w} + s_{wq}(.)^2(\check{q}_a + \Delta\bar{q}^2)\Delta\check{w} \\ & + 2s_{bq}(.)s_{hq}(.)\check{q}_a\Delta\bar{b}\Delta\bar{h} + s_{qq}(.)s_{bq}(.)\check{q}_a\Delta\bar{q}\Delta\bar{b} \\ & + s_{qq}(.)s_{hq}(.)\check{q}_a\Delta\bar{q}\Delta\bar{h} + 2s_b(.)s_{bq}(.)\check{b}_a\Delta\bar{q} \\ & + 2s_h(.)s_{hq}(.)\check{h}_a\Delta\bar{q} + 2s_w(.)s_{hw}(.)\Delta\check{w}\Delta\bar{h} \\ & + 2s_w(.)s_{qw}(.)\Delta\check{w}\Delta\bar{q} + 2s_q(.)s_{qq}(.)\check{q}_a\Delta\bar{q} \\ & + 2s_q(.)s_{bq}(.)\check{q}_a\Delta\bar{b} + 2s_q(.)s_{qh}(.)\check{q}_a\Delta\bar{h}\end{aligned}$$

Because there is no Dependence Approximation:

$$\Delta\bar{s} = (\bar{s}_a - \bar{s}_x) + (\bar{s}_x - \bar{s}_m) = (\bar{s}_x - \bar{s}_m)$$

$$\Delta\check{s} = (\check{s}_a - \check{s}_x) + (\check{s}_x - \check{s}_m) = (\check{s}_x - \check{s}_m)$$

For $p = 160$, these correction terms are:

$$\Delta\bar{s} = -.472$$

$$\Delta\check{s} = 5.797$$

The modeled mean and variance of s for $p = 160$ are:

$$\bar{s}_m = 7.014$$

$$\check{s}_m = 11.317$$

So, the estimated authentic mean and variance of profit s for $p = 160$ are:

$$\bar{s}_a = \bar{s}_m + \Delta\bar{s} = 6.542$$

$$\check{s}_a = \check{s}_m + \Delta\check{s} = 17.114$$

We see that, for $p = 160$, the analysis overstates the expected profit by about 7.2% and understates the variance by about 34% and the standard deviation by about 19%.

Cumulative effect of the modeling approximations on the certain equivalent:

By treating the utility function as a submodel, we can determine the cumulative effect of all the modeling approximations on the certain equivalent.

The utility function is:

$$u(s) = 1 - e^{-.1\,s}$$

Taking derivatives:

$$u'(s) = .1\,e^{-.1\,s} \qquad u''(s) = -.01\,e^{-.1\,s}$$

For $p = 160$, these derivatives evaluated at $\bar{s}_m$ are:

$$u'(\bar{s}_m) = .0496$$

$$u''(\bar{s}_m) = -4.96 \times 10^{-3}$$

The correction term for the effect on expected utility of the Distribution Approximation from (2.11) is:

$$\Delta\bar{u} = u'(\bar{s}_m)\Delta\bar{s} + u''(\bar{s}_m)\,(\Delta\check{s} + \Delta\bar{s}^2)$$

For $p = 160$, this correction term is:

$$\Delta\bar{u} = -.0383$$

The modeled expected utility for $p = 160$ is:

$$\bar{u}_m = .4764$$

So the estimated authentic expected utility for $p = 160$ is:

$$\bar{u}_a = \bar{u}_m + \Delta\bar{u} = .4381$$

This corresponds to a certain equivalent of:

$$CE_a = u^{-1}(\bar{u}_a) = 5.76$$

The modeled certain equivalent for $p = 160$ is:

$$CE_m = 6.47$$

We see that, for $p = 160$, the modeling approximations have caused the analysis to overstate the certain equivalent by about $710,000 or about 12%.

By repeating this procedure for different values of p, we find the following:

p	Modeled CE_m	Est. authentic CE_a	Overstatement by analysis
150	5.87	5.20	13%
160	6.47	5.76	12%
170	6.40	5.67	13%

After taking into account the effect of the modeling approximations, we see that the optimal decision is still $p = 160$.

Effect of the modeling approximations on the value of clairvoyance:

If we apply the methodology to the probabilistic sensitivity analysis, we can estimate the effect of the modeling approximations on the value of clairvoyance. We find as a result that the analysis slightly overstates the value of clairvoyance:

Variable	Value of Clairvoyance	
	Modeled	Est. authentic
Competitor's price: y	$378,000	$327,000
Economic indicator: f	0	0
Raw material cost: h	$ 46,000	$ 43,000
Other var. costs: b	$ 31,000	$ 25,000

Summary

We see that the methodology presented in Chapter 2 can be used to determine the effect on the profit lottery and on the certain equivalent of the modeling approximations made in the analysis. The methodology thus provides us with an indication of the adequacy of the modeling in a decision analysis.

In this example, the modeling appears to be quite adequate, for the methodology indicates that the modeling approximations do not affect the optimal decision.

CHAPTER 4

STOCHASTIC MODELS

4.1 Introduction

In the preceding chapters, we have dealt solely with deterministic models -- those models that yield a unique value of the outcome variable for fixed values of the state and decision variables. We can now look at stochastic models -- those models that yield a probability distribution on the outcome variable for fixed values of the state and decision variables.

We use stochastic models to represent situations in which we are uncertain about the dependence of the outcome variable on the state variables but cannot conveniently attribute this uncertainty to additional state variables. Indeed, in these situations, we often view the dependence relationship as being inherently uncertain. Generally, we possess some data (e.g., long-run frequencies) about the uncertain relationship to help us define the stochastic model.

For example, we use the Bernoulli stochastic model to express our uncertainty about the number of heads obtained in n tosses of a coin because we find it easier to think of the outcome of a coin toss as being inherently uncertain than to think of it as being dependent on such extraneous factors as rotational forces and surface characteristics.

In this chapter, we shall extend the methodology developed in Chapter 2 for deterministic models to stochastic models. The methodology is useful only for relatively simple stochastic models, such as Bernoulli or small Markov models.

4.2 Extension of the Methodology

In a decision analysis, we ideally want the authentic profit lottery:

$$\{v|d,\mathcal{S}\} = \int_{\underline{s}} \{v|d,\underline{s},\mathcal{S}\}\{\underline{s}|\mathcal{S}\}$$

Suppose, however, that we use a stochastic model to compute the profit lottery, as follows:

$$\{v|d,\mathcal{S}\}_m = \sum_{\underline{s}} \{v|d,\underline{s},\mathcal{S}\}_m \; p(\underline{s})$$

where the stochastic model is specified by the conditional probability distribution $\{v|d,\underline{s},\mathcal{S}\}_m$. We see that the discrepancy between the modeled and authentic profit lotteries is due to the two modeling approximations:

1. The approximation of the probability distribution on the state variable vector $\underline{s}$ by the discrete probability function p (Distribution Approximation):

$$(4.1) \qquad p(\underline{s}) \simeq \{\underline{s}|\mathcal{S}\}$$

2. The approximation of the conditional probability distribution on v given d and $\underline{s}$ by the stochastic model (Dependence Approximation):

$$(4.2) \qquad \{v|d,\underline{s},\mathcal{S}\}_m \approx \{v|d,\underline{s},\mathcal{S}\}$$

We want to quantitatively characterize these modeling approximations and determine their effect on the profit lottery.

More specifically, letting v_a denote the authentic outcome variable and v_y denote the outcome variable of the stochastic model, we want to determine the correction terms:

$$\Delta\bar{v} = (\bar{v}_a - \bar{v}_y)$$

$$\Delta\check{v} = (\check{v}_a - \check{v}_y)$$

Let us first define:

$$\left.\begin{aligned} g(d,\underline{s}) &\equiv \langle v_y|d,\underline{s},\mathcal{S}\rangle \\ f(d,\underline{s}) &\equiv \check{\langle} v_y|d,\underline{s},\mathcal{S}\rangle \end{aligned}\right\} \text{ the conditional mean and variance of the stochastic model}$$

Then, we can write:

$$(4.3) \qquad v_y = g(d,\underline{s}) + \varepsilon$$

where ε is a random variable whose conditional mean and variance are:

$$\langle \varepsilon | d,\underline{s},\delta \rangle = \langle v_y | d,\underline{s},\delta \rangle - g(d,\underline{s}) = 0$$

$$^{v}\langle \varepsilon | d,\underline{s},\delta \rangle = {}^{v}\langle v_y | d,\underline{s},\delta \rangle = f(d,\underline{s})$$

Then, using $g(d,\underline{s})$ as a reference deterministic model, we can write the Dependence Approximation in (4.2) as two separate approximations:

$$\{v_y | d,\underline{s},\delta\}_m \approx \delta(v - g(d,\underline{s})) \approx \{v_a | d,\underline{s},\delta\}$$

We characterize the first of these approximations by the random variable ε as in (4.3) above and the second, as in Chapter 2, by the random variable e :

(4.4) $$v_a = g(d,\underline{s}) + e$$

where the conditional mean $\langle e | d,\underline{s},\delta \rangle$ and variance $^{v}\langle e | d,\underline{s},\delta \rangle$ are directly assessed.

Next, employing the following notation from Chapter 2:

$\underline{n}$ = the state variable vector whose distribution is $p(\underline{s})$

$\underline{x}$ = the state variable vector whose distribution is $\{\underline{s} | \delta\}$

we characterize the Distribution Approximation in (4.1) by the correction terms:

$$\Delta\bar{\underline{s}} = \bar{\underline{x}} - \bar{\underline{n}}$$

$$\Delta\check{\underline{s}} = \check{\underline{x}} - \check{\underline{n}}$$

We use four different subscripts for the outcome variable (see Figure 4.1):

v_y = the outcome variable of the stochastic model

v_m = the outcome variable of the reference deterministic model

v_x = the outcome variable corrected for the Distribution Approximation

v_a = the authentic outcome variable

Then, we can write:

$$v_m = g(d,\underline{n})$$

$$v_x = g(d,\underline{x})$$

$$v_y = g(d,\underline{n}) + \varepsilon = v_m + \varepsilon \qquad (4.5)$$

$$v_a = g(d,\underline{x}) + e = v_x + e \qquad (4.6)$$

Taking the mean and variance of (4.5), we have:

Figure 4.1: Definition of Subscripted Outcome Variables

v_y = Outcome variable of stochastic model

$$\{v_y | d,\&\} = \sum_{\underline{s}} \{v | d,\underline{s},\&\}_m \; p(\underline{s})$$

↓

> Correction for:
>
> $\{v | d,\underline{s},\&\}_m \approx \delta(v - g(d,\underline{s}))$

↓

v_m = Outcome variable of reference deterministic model

$$\{v_m | d,\&\} = \sum_{\underline{s}} \delta(v - g(d,\underline{s})) \; p(\underline{s})$$

↓

> Correction for:
>
> $p(\underline{s}) \approx \{\underline{s} | \&\}$

↓

v_x = Outcome variable corrected for Distribution Approximation

$$\{v_x | d,\&\} = \int_{\underline{s}} \delta(v - g(d,\underline{s}))\{\underline{s} | \&\}$$

↓

> Correction for:
>
> $\delta(v - g(d,\underline{s})) \approx \{v | d,\underline{s},\&\}$

↓

v_a = Authentic outcome variable

$$\{v_a | d,\&\} = \int_{\underline{s}} \{v | d,\underline{s},\&\}\{\underline{s} | \&\}$$

$$\bar{v}_y = \bar{v}_m + \bar{\varepsilon}$$

$$\check{v}_y = \check{v}_m + \check{\varepsilon} + 2\mathrm{cov}(v_m,\varepsilon)$$

But, because $\langle \varepsilon | d,\underline{s},\delta \rangle = 0$, $\bar{\varepsilon} = 0$ and $\mathrm{cov}(v_m,\varepsilon) = 0$.

So,

(4.7) $$\bar{v}_y = \bar{v}_m$$

(4.8) $$\check{v}_y = \check{v}_m + \check{\varepsilon}$$

where:

$$\check{\varepsilon} = \sum_{\underline{s}} \check{\langle} \varepsilon | d,\underline{s},\delta \rangle \; p(\underline{s})$$

(4.9)

$$= \sum_{\underline{s}} f(d,\underline{s}) \; p(\underline{s})$$

Next, taking the mean and variance of (4.6), we have:

(4.10) $$\bar{v}_a = \bar{v}_x + \bar{e}$$

(4.11) $$\check{v}_a = \check{v}_x + \check{e} + 2\mathrm{cov}(v_x,e)$$

where (see Appendix A.4) :

$$(4.12) \qquad \bar{e} = \int_{\underline{s}} \langle e|d,\underline{s},\delta\rangle\{\underline{s}|\delta\}$$

$$(4.13) \qquad \check{e} = \int_{\underline{s}} \left({}^{\vee}\langle e|d,\underline{s},\delta\rangle + \langle e|d,\underline{s},\delta\rangle^2\right)\{\underline{s}|\delta\} \quad - \bar{e}^2$$

$$(4.14) \qquad \mathrm{cov}(v_x,e) = \int_{\underline{s}} g(d,\underline{s})\langle e|d,\underline{s},\delta\rangle\{\underline{s}|\delta\} \quad - \bar{v}_x\bar{e}$$

Then, subtracting (4.7) from (4.9) and (4.8) from (4.10), we have:

$$(4.15) \qquad \Delta\bar{v} = (\bar{v}_a - \bar{v}_y) = (\bar{v}_x - \bar{v}_m) + \bar{e}$$

$$(4.16) \qquad \Delta\check{v} = (\check{v}_a - \check{v}_y) = (\check{v}_x - \check{v}_m) + (\check{e} - \check{\varepsilon}) + 2\mathrm{cov}(v_x,e)$$

where $\check{\varepsilon}$, $\bar{e}$, $\check{e}$ and $\mathrm{cov}(v_x,e)$ are calculated as in (4.9) and (4.12) through (4.14) above and where $(\bar{v}_x - \bar{v}_m)$ and $(\check{v}_x - \check{v}_m)$ are the correction terms for the effect of the Distribution Approximation, given by (2.11) and (2.12) for the single state variable case and by (2.27) and (2.28) for the multivariate case. The equations (4.15) and (4.16) above yield the desired correction terms for the outcome variable.

As a special case, suppose that the stochastic model fully expresses our uncertainty about the dependence relationship:

$$\{v_y|d,\underline{s},\delta\}_m = \{v_a|d,\underline{s},\delta\}$$

Then,

$$\langle v_a | d,\underline{s},\delta\rangle = \langle v_y | d,\underline{s},\delta\rangle$$

so,

$$\langle e | d,\underline{s},\delta\rangle = \langle v_a | d,\underline{s},\delta\rangle - g(d,\underline{s}) = 0$$

and, therefore,

$$\bar{e} = 0 \quad \text{and} \quad \text{cov}(v_x, e) = 0$$

Also,

$$\check{\langle} e | d,\underline{s},\delta\rangle = \check{\langle} v_y | d,\underline{s},\delta\rangle = f(d,\underline{s})$$

and, therefore, from (4.13),

$$\check{e} = \int_{\underline{s}} f(d,\underline{s})\{\underline{s}|\delta\} \tag{4.17}$$

Then, from (4.9) and (4.17),

$$(\check{e} - \check{\varepsilon}) = \int_{\underline{s}} f(d,\underline{s})\{\underline{s}|\delta\} - \sum_{\underline{s}} f(d,\underline{s})\, p(\underline{s})$$

From (2.33), we see that this can be estimated as:

$$(\check{e} - \check{\varepsilon}) \simeq \sum_i f_i(d,\bar{\underline{n}})\Delta\bar{s}_i + \tfrac{1}{2}\sum_i f_{ii}(d,\bar{\underline{n}})\Delta\check{s}_i \tag{4.18}$$
$$\tfrac{1}{2}\sum_i \sum_j f_{ij}(d,\bar{\underline{n}})\Delta\bar{s}_i\Delta\bar{s}_j$$

So, for this special case, the correction terms in (4.15) and (4.16) become:

$$(4.19) \qquad \Delta\bar{v} = (\bar{v}_x - \bar{v}_m)$$

$$(4.20) \qquad \Delta\check{v} = (\check{v}_x - \check{v}_m) + (\check{e} - \check{\varepsilon})$$

4.3 Example: A Game

Consider the following game: We draw one ball from an urn containing 25 balls that are numbered consecutively from 1 to 25. Letting n be the number on the drawn ball, we then flip a thumbtack n times. Finally, letting r be the number of "heads" (point up) obtained, we receive a prize of v dollars, where $v = r^2$.

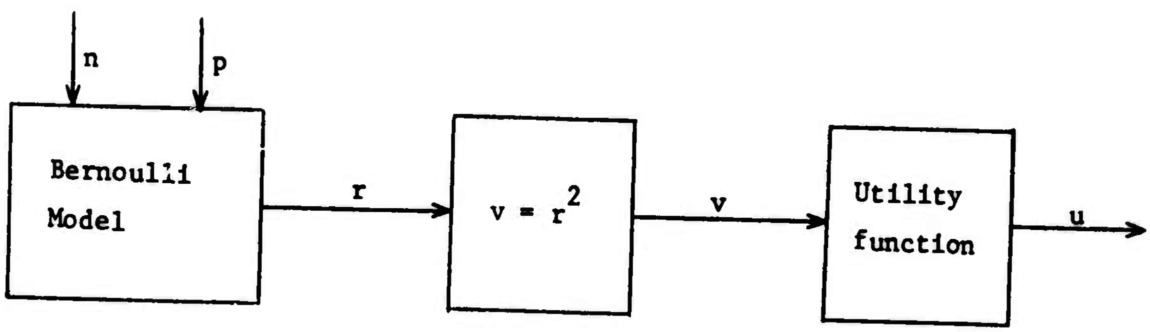
Should we pay \$25 to play this game? Our utility function on dollars is exponential with risk aversion coefficient $\gamma = .04$:

$$u(v) = 1 - e^{-.04v}$$

We use the Bernoulli stochastic model to analyze this game. (See Figure 4.2) The state variables are n , the number of flips, and p , the long-run frequency of "heads" in many flips of the thumbtack. Given fixed values of these state variables, the Bernoulli model yields the following probability distribution on the number of "heads" r :

$$\{r|n,p,\mathscr{s}\}_m = \binom{n}{r}(p)^r(1 - p)^{n-r}$$

Figure 4.2: Analysis of the Game



n = number on drawn ball

p = long-run frequency of heads

r = number of heads in n tosses

v = monetary payoff

u = utility

For the purposes of analysis, we use the following distributions on the state variables:

$$\{n|\mathcal{s}\}_m = \begin{cases} \frac{1}{25} & n = 1,2,\ldots 25 \\ 0 & \text{otherwise} \end{cases}$$

$$\{p|\mathcal{s}\}_m = \delta(p - p_0) \quad \text{where } p_0 = .5$$

Using a computer, we calculate the resulting probability distribution on r :

$$\{r|\mathcal{s}\}_m = \sum_n \sum_p \{r|n,p,\mathcal{s}\}_m \{n|\mathcal{s}\}_m \{p|\mathcal{s}\}_m$$

$$= \frac{1}{25} \sum_n \binom{n}{r} (.5)^n$$

Next, we calculate the corresponding distribution on v :

$$\{v|\mathcal{s}\}_m = \sum_r \{v|r,\mathcal{s}\}\{r|\mathcal{s}\}_m$$

$$= \sum_r \delta(v - r^2)\{r|\mathcal{s}\}_m$$

Finally, we compute the expected utility and the certain equivalent of the game:

$$\bar{u}_m = \sum_v \{v|\mathcal{s}\}_m \, u(v)$$

$$CE_m = u^{-1}(\bar{u}_m)$$

We find that:

$$\bar{r}_m = 6.50 \qquad \check{r}_m = 16.25$$

$$\bar{v}_m = 58.50 \qquad \check{v}_m = 3586.05$$

$$\bar{u}_m = 0.6469$$

$$CE_m = 26.03$$

The analysis indicates that we should be willing to pay up to \$26.03 to play this game.

We can now ask: What modeling approximations have we made in this analysis and what effect do they have on its results? In this case, there is just one approximation -- the use of a fixed value $p_0 = .5$ for the long-run fraction of "heads" instead of a probability distribution on p. Suppose that we assess the following mean and variance for p :

$$\bar{p} = p_0 = .5$$

$$\check{p} = .01$$

Then, we characterize the approximation by the correction terms:

$$\Delta\bar{p} = 0$$

$$\Delta\check{p} = .01$$

We define as the reference deterministic model the conditional mean of the Bernoulli model, letting:

$$g(n,p) = \langle r \mid n,p,\delta \rangle = n\,p$$

$$f(n,p) = \check{\langle} r \mid n,p,\delta \rangle = n\,p\,(1-p)$$

Taking partial derivatives of $g(n,p)$:

$$g_n = p \qquad g_{nn} = 0$$

$$g_p = n \qquad g_{pp} = 0$$

$$g_{np} = 1$$

Then, noting that $\Delta\bar{n}$, $\Delta\check{n}$ and $\Delta\bar{p}$ are all zero, we have from (2.27) and (2.28) :

$$(\bar{r}_x - \bar{r}_m) = 0$$

$$\begin{aligned}(\check{r}_x - \check{r}_m) &= g_p(\bar{n},\bar{p})^2 \Delta\check{p} + g_{np}(\bar{n},\bar{p})^2 (\check{p}_a \check{n} - \check{p}_m \check{n}) \\ &= \bar{n}^2 \Delta\check{p} + \check{n}\Delta\check{p} \\ &= (\bar{n}^2 + \check{n})\Delta\check{p}\end{aligned}$$

Because the Bernoulli model fully expresses our uncertainty about the number of "heads" given n and p , we can use the results of Section 4.2 for the special case. Taking partial derivatives of $f(n,p)$ with respect to p :

$$f_p = n - 2np$$

$$f_{pp} = -2n$$

Then, from (4.18), we have:

$$(\check{e} - \check{\varepsilon}) = \tfrac{1}{2} f_{pp}(\bar{n},\bar{p})\Delta\check{p}$$

$$= -\bar{n}\Delta\check{p}$$

So, from (4.19) and (4.20), we have:

$$\Delta\bar{r} = (\bar{r}_x - \bar{r}_m) = 0$$

$$\Delta\check{r} = (\check{r}_x - \check{r}_m) + (\check{e} - \check{\varepsilon}) = (\bar{n}^2 + \check{n} - \bar{n})\Delta\check{p}$$

Substituting, we find that:

$\Delta\bar{r}$	=	0	$\bar{r}_a$	=	6.50
$\Delta\check{r}$	=	2.08	$\check{r}_a$	=	18.33

So, the fixation of p has no effect on the expected value of r but causes the analysis to understate its variance by about 11%.

We can now determine the effect of the approximation on our winnings v. Taking derivatives of $v = r^2$:

$$v' = 2r$$

$$v'' = 2$$

Then, noting that $\Delta\bar{r} = 0$, we have from (2.11) and (2.12):

$$\Delta\bar{v} = \tfrac{1}{2}v''(\bar{r}_m)\Delta\check{r}$$

$$= \Delta\check{r}$$

$$\Delta\check{v} = v'(\bar{r}_m)^2\Delta\check{r} + \tfrac{1}{2}v''(\bar{r}_m)^2 \Delta\check{r}(\check{r}_a + \check{r}_m)$$

$$= 4\bar{r}_m^2\Delta\check{r} + 2\Delta\check{r}(2\check{r}_m + \Delta\check{r})$$

Substituting, we find that:

$$\Delta\bar{v} = 2.08 \qquad \bar{v}_a = 60.58$$

$$\Delta\check{v} = 495.37 \qquad \check{v}_a = 4081.42$$

We see that the fixation of p causes the analysis to understate both the expected value and the variance of v by about 3.4% and 12% , respectively.

Finally, we can determine the effect of the approximation on the expected utility and the certain equivalent. Taking derivatives of the utility function:

$$u(v) = 1 - e^{-.04v}$$

$$u' = .04e^{-.04v}$$

$$u'' = -.0016\, e^{-.04v}$$

Then, from (2.26), we have:

$$\Delta\bar{u} = u'(\bar{v}_m)\Delta\bar{v} + \tfrac{1}{2}u''(\bar{v}_m)(\Delta\check{v} + \Delta\bar{v}^2)$$

$$= .04\, e^{-.04\bar{v}_m}\,(\Delta\bar{v} - .02(\Delta\check{v} + \Delta\bar{v}^2))$$

Substituting, we find that:

$$\Delta\bar{u} = -0.0305 \qquad \bar{u}_a = 0.6164$$

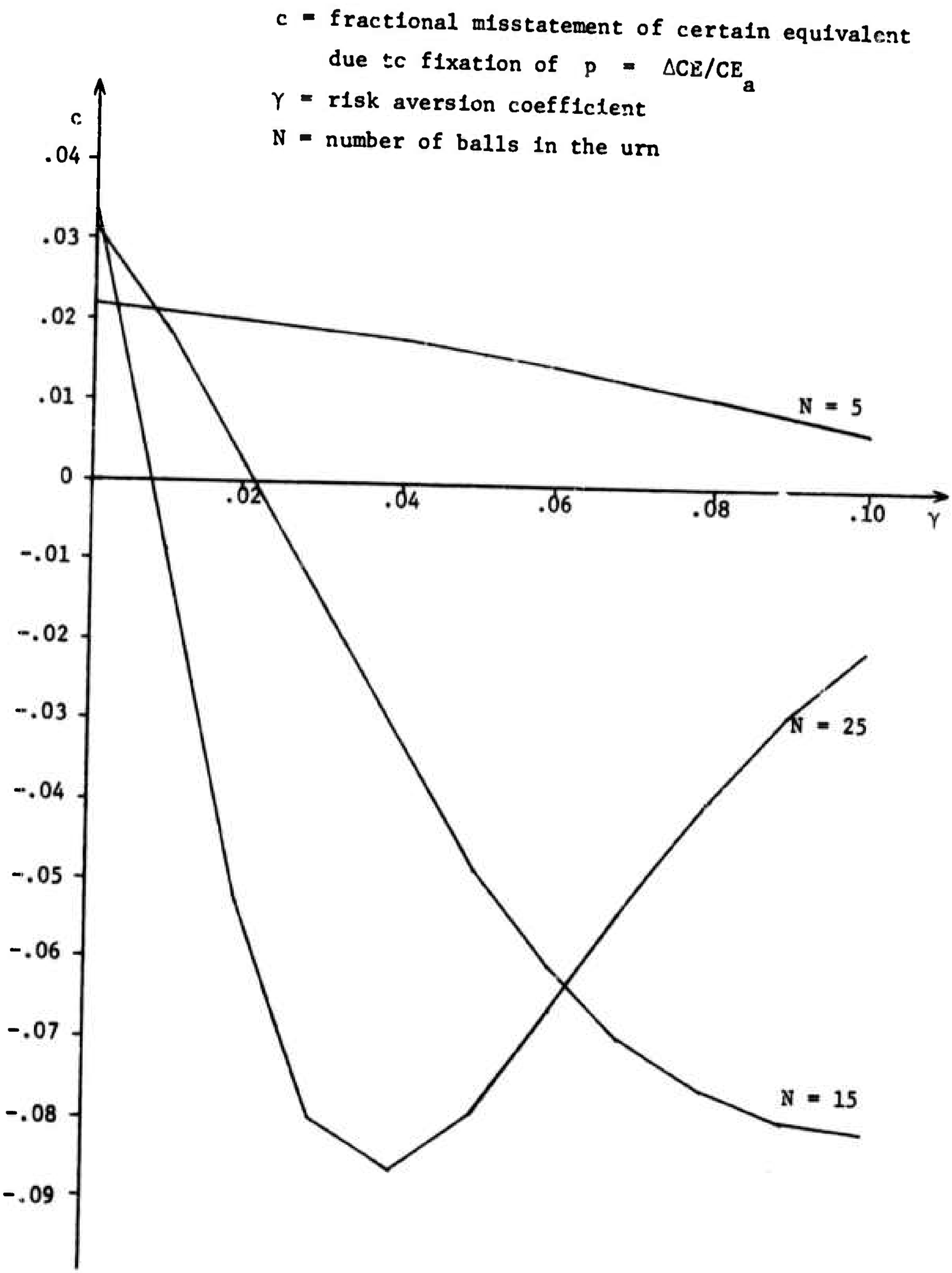
And:

$$CE_a = -25\ \ln(1 - \bar{u}_a) = \$23.96$$

So, we see that the fixation of p causes the analysis to overstate the certain equivalent by about 8.6% and that we should not be willing to pay $25 to play the game.

It is interesting to see how the misstatement of the certain equivalent caused by the fixation of p varies with the degree of our risk aversion. Figure 4.3 shows the percentage misstatement of the certain equivalent as a function of our risk aversion coefficient γ for the game played with 5, 15 and 25 balls in the urn. Note that the misstatement of the certain equivalent may be either positive or negative, depending on our risk preference, and that the degree of misstatement does not necessarily vary monotonically with the risk aversion coefficient.

Figure 4.3: Risk Sensitivity of Misstatement of CE



CONCLUSION

As we have seen, the role of modeling in decision analysis is to help us obtain the authentic profit lottery by simplifying our assessment task. We think of the models in decision analysis, then, as subjective expressions of our uncertain understanding of the world rather than as objective descriptions of reality.

We can now answer the questions about modeling that we raised in the Introduction.

First, what is "goodness" in a model and can we quantify it? Because we think of a model as an approximate expression of our uncertainty and not as a description of the real-world, an index of realism would not provide a meaningful measure of goodness. Rather, the measure of goodness that we want is how well the model represents our uncertainty about the dependence relationship between outcome and state variables. We can quantify this measure of goodness with the random variable e :

$$v = g(d,\underline{s}) + e$$

Then, we can use the methodology of Chapter 2 to show how the goodness of the model affects the results of the decision analysis by determining its contribution to the correction terms on the profit lottery $\bar{\Delta v}$ and $\check{\Delta v}$.

Next, how should we choose among alternative models? Using the measure of goodness defined above, we should choose among competing models according to how "close" they get us to the authentic profit lottery. When our state of information about the decision environment includes a large amount of data, we can use a probability-space of models to help us choose among several alternative model forms, being careful that we meaningfully interpret the probability assigned to each model in the space. (See Section 1.6.)

Finally, how should we decide when to do more modeling? The reason that we use modeling in decision analysis at all is that we judge the directly assessed profit lottery to be so "far" from the authentic profit lottery that it is worthwhile spending the time and effort to reduce our secondary uncertainty. Likewise, our decision to do more modeling should depend on whether or not it is worth the additional effort to get even "closer" to the authentic profit lottery.

We use the concept of the value of perfect information to help us decide when to collect more data. The analogous concept that we need to help us decide when to do more modeling is not the value of perfect modeling but rather the value of obtaining the authentic profit lottery. This value is the most we should pay to completely resolve our secondary uncertainty through further modeling. Unfortunately, it is exceedingly difficult to exercise this concept, since to do so, we must assess the joint distribution on the authentic certain equivalents for all decision alternatives.

We can, however, get an indication of the value of additional modeling by using the methodology of Chapter 2 to estimate the difference between the modeled and authentic profit lotteries. The greater the difference, the greater the presumed value of additional modeling. Furthermore, we can use the methodology to indicate where more modeling would be most effective by showing the contribution of each part of the model to the discrepancy in the profit lotteries.

APPENDICES

Appendix A.1: Effect of the Distribution Approximation

This appendix supplements Section 2.2, pp. 39-40.

For notational ease, let:

$$g(.) = g(d,\bar{n})$$

$$\overset{3}{s} = \overline{(s - \bar{s})^3} \quad \text{, third central moment}$$

$$\overset{4}{s} = \overline{(s - \bar{s})^4} \quad \text{, fourth central moment}$$

We characterize the Distribution Approximation by:

$$\Delta\bar{s} = \bar{x} - \bar{n}$$

$$\Delta\check{s} = \check{x} - \check{n}$$

We expand $v_m = g(d,n)$ about $\bar{n}$:

(a.1) $$v_m \simeq g(.) + g'(.)\ (n - \bar{n}) + \tfrac{1}{2}g''(.)\ (n - \bar{n})^2$$

Taking the mean:

(a.2) $$\bar{v}_m \simeq g(.) + \tfrac{1}{2}g''(.)\ \check{n}$$

Squaring (a.1) and taking the mean:

$$v_m^2 \simeq g(.)^2 + g'(.)^2\ (n - \bar{n})^2 + \tfrac{1}{4}g''(.)^2\ (n - \bar{n})^4$$
$$+ 2g(.)g'(.)\ (n - \bar{n}) + g(.)g''(.)\ (n - \bar{n})^2$$
$$+ g'(.)g''(.)\ (n - \bar{n})^3$$

(a.3) $$\overline{v_m^2} \simeq g(.)^2 + g'(.)^2 \overset{\vee}{n} + \tfrac{1}{4} g''(.)^2 \overset{4}{n}$$

$$+ g(.)g''(.)\overset{\vee}{n} + g'(.)g''(.)\overset{3}{n}$$

Squaring (a.2) :

(a.4) $$\bar{v}_m^{\,2} \simeq g(.)^2 + \tfrac{1}{4} g''(.)^2 \overset{\vee}{n}{}^2 + g(.)g''(.)\overset{\vee}{n}$$

Subtracting (a.4) from (a.3) :

(a.5) $$\overset{\vee}{v}_m \simeq g'(.)^2 \overset{\vee}{n} + \tfrac{1}{4} g''(.)^2 (\overset{4}{n} - \overset{\vee}{n}{}^2) + g'(.)g''(.)\overset{3}{n}$$

To simplify this expression, we assume the following (see Appendix A.2) :

$$\overset{3}{n} = 0$$

$$\overset{4}{n} = 3\,\overset{\vee}{n}{}^2$$

Then (a.5) becomes:

(a.6) $$\overset{\vee}{v}_m \simeq g'(.)^2 \overset{\vee}{n} + \tfrac{1}{2} g''(.)^2 \overset{\vee}{n}{}^2$$

Similarly, we expand $v_x = g(d,x)$ about $\bar{n}$:

(a.7) $$v_x \simeq g(.) + g'(.)\,(x - \bar{n}) + \tfrac{1}{2} g''(.)\,(x - \bar{n})^2$$

Taking the mean:

(a.8) $$\bar{v}_x \simeq g(.) + g'(.)\,(\bar{x} - \bar{n}) + \tfrac{1}{2} g''(.)\,\overline{(x - \bar{n})^2}$$

Squaring (a.7) and taking the mean:

$$v_x^2 \simeq g(.)^2 + g'(.)^2 (x - \bar{n})^2 + \tfrac{1}{4}g''(.)^2 (x - \bar{n})^4$$

$$+ 2g(.)g'(.) (x - \bar{n}) + g(.)g''(.) (x - \bar{n})^2$$

$$+ g'(.)g''(.) (x - \bar{n})^3$$

(a.9) $$\overline{v_x^2} \simeq g(.)^2 + g'(.)^2 \overline{(x - \bar{n})^2} + \tfrac{1}{4}g''(.)^2 \overline{(x - \bar{n})^4}$$

$$+ 2g(.)g'(.) (\bar{x} - \bar{n}) + g(.)g''(.) \overline{(x - \bar{n})^2}$$

$$+ g'(.)g''(.) \overline{(x - \bar{n})^3}$$

Squaring (a.8) :

(a.10) $$\bar{v}_x^2 \simeq g(.)^2 + g'(.)^2 (\bar{x} - \bar{n})^2 + \tfrac{1}{4}g''(.)^2 \overline{(x - \bar{n})^2}^2$$

$$+ 2g(.)g'(.) (\bar{x} - \bar{n}) + g(.)g''(.) \overline{(x - \bar{n})^2}$$

$$+ g'(.)g''(.) (\bar{x} - \bar{n}) \overline{(x - \bar{n})^2}$$

Subtracting (a.10) from (a.9)

(a.11) $$\check{v}_x \simeq g'(.)^2\left[\overline{(x - \bar{n})^2} - (\bar{x} - \bar{n})^2\right]$$

$$+ \tfrac{1}{4}g''(.)^2 \left[\overline{(x - \bar{n})^4} - \overline{(x - \bar{n})^2}^2\right]$$

$$+ g'(.)g''(.)\left[\overline{(x - \bar{n})^3} - (\bar{x} - \bar{n}) \overline{(x - \bar{n})^2}\right]$$

Note that:

$$(\bar{x} - \bar{n}) = \overline{\Delta s}$$

$$\overline{(x - \bar{n})^2} = \overline{((x - \bar{x}) + (\bar{x} - \bar{n}))^2} = \check{x} + \overline{\Delta s}^2$$

$$\overline{(x - \bar{n})^3} = \overline{((x - \bar{x}) + (\bar{x} - \bar{n}))^3} = \overset{3}{x} + 3\check{x}\overline{\Delta s} + \overline{\Delta s}^3$$

$$\overline{(x - \bar{n})^4} = \overline{((x - \bar{x}) + (\bar{x} - \bar{n}))^4} = \overset{4}{x} + 4\overset{3}{x}\overline{\Delta s} + 6\check{x}\overline{\Delta s}^2 + \overline{\Delta s}^4$$

Substituting into (a.8) and (a.11) :

(a.12) $$\bar{v}_x \simeq g(.) + g'(.)\overline{\Delta s} + \tfrac{1}{2}g''(.)\,(\check{x} + \overline{\Delta s}^2)$$

(a.13) $$\check{v}_x \simeq g'(.)^2\,\check{x} + \tfrac{1}{4}g''(.)^2(\overset{4}{x} - \check{x}^2 + 4\overset{3}{x}\overline{\Delta s} + 4\,\check{x}\overline{\Delta s}^2)$$

$$+ g'(.)g''(.)\,(\overset{3}{x} + 2\check{x}\overline{\Delta s})$$

Again, we use the simplifying assumptions (see Appendix A.2) :

$$\overset{3}{x} = 0$$

$$\overset{4}{x} = 3\,\check{x}^2$$

Then, (a.13) becomes:

(a.14) $$\check{v}_x \simeq g'(.)^2\,\check{x} + \tfrac{1}{2}g''(.)^2(\check{x}^2 + 2\check{x}\overline{\Delta s}^2)$$

$$+ 2g'(.)g''(.)\,\check{x}\overline{\Delta s}$$

In summary, we have:

$$\bar{v}_m \approx g(.) + \tfrac{1}{2}g''(.)\,\check{n} \qquad \text{(a.2)}$$

$$\check{v}_m \approx g'(.)^2\,\check{n} + \tfrac{1}{2}g''(.)^2\,\check{n}^2 \qquad \text{(a.6)}$$

$$\bar{v}_x \approx g(.) + g'(.)\Delta\bar{s} + \tfrac{1}{2}g''(.)\,(\check{x} + \Delta\bar{s}^2) \qquad \text{(a.12)}$$

$$\check{v}_x \approx g'(.)^2\,\check{x} + \tfrac{1}{2}g''(.)^2\,(\check{x}^2 + 2\check{x}\Delta\bar{s}^2) + 2g'(.)g''(.)\,\check{x}\Delta\bar{s} \qquad \text{(a.14)}$$

These equations are shown as Equations (2.7) through (2.10). By subtraction, we get:

$$(\bar{v}_x - \bar{v}_m) \approx g'(.)\Delta\bar{s} + \tfrac{1}{2}g''(.)\,(\Delta\check{s} + \Delta\bar{s}^2)$$

$$(\check{v}_x - \check{v}_m) \approx g'(.)^2\Delta\check{s} + \tfrac{1}{2}g''(.)^2\left[\Delta\check{s}(\check{x} + \check{n}) + 2\check{x}\Delta\bar{s}^2\right] + 2g'(.)g''(.)\,\check{x}\Delta\bar{s}$$

These equations yield the correction terms for the effect of the Distribution Approximation and are shown as Equations (2.11) and (2.12).

Appendix A.2: Simplifying Assumptions

In Chapter 2, we assessed correction terms for the mean and variance of the state variable and calculated the resulting correction terms for the mean and variance of the outcome variable. For several linked submodels, we saw that the output variable of one submodel becomes the input variable of the next and that we can sequentially calculate the correction terms for each intermediate variable.

Clearly, since we want to deal only with the mean and variance of each variable, we must eliminate all terms of higher order. In developing the correction terms for the effect of the Distribution Approximation in Appendix A.1, we encountered the third and fourth central moments of the input variable and eliminated them by assuming the following:

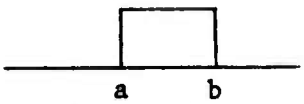
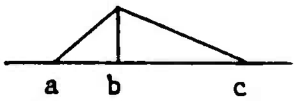
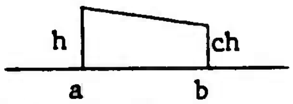
$$\overline{(s - \bar{s})^3} = 0$$

$$\overline{(s - \bar{s})^4} = 3\,\check{s}^2$$

If we know only the mean and variance of a variable, we cannot reliably estimate its third central moment, which is a measure of the skewness of the distribution about the mean. We cannot even predict its sign. Therefore, in the methodology, I make the assumption that the third central moment is equal to zero, which is true if the distribution is symmetric.

On the other hand, we can estimate the fourth central moment if we know the variance; both are measures of the distribution's dispersion about the mean. In fact, it turns out that the fourth central moment is proportional, or nearly so, to the square of the variance for many of the named distributions. (See Table A.1) The ratio of the fourth central moment to the variance squared ranges from 1.8 for the uniform distribution to 9 for the exponential, but many of the distributions have ratios of about 3. Therefore, I assume for the purposes of the methodology that the fourth central moment is equal to three times the variance squared.

Table A.1

	Distribution (parameters)	Ratio of the fourth central moment to the variance squared	
1.	Uniform (a,b) 	$\frac{9}{5}$	for any a,b
2.	Triangular (a,b,c) 	$\frac{12}{5}$	for any a,b,c
3.	Trapezoidal (a,b,c) 	$\frac{12}{5}\,\frac{(1+8c+9c^2+8c^3+c^4)}{(1+8c+18c^2+8c^3+c^4)}$	for any a,b
4.	Normal (μ,σ)	3	for any μ,σ
5.	Beta (a,b)	$3\,\frac{(b+1)}{(b+2)(b+3)}\left[(b-6)-\frac{2b^2}{a(a-b)}\right]$	
6.	Beta with $b = 2a$	$3\,\frac{(b+1)}{(b+3)}$	
7.	Exponential (λ)	9	for any λ
8.	Laplace (λ)	6	for any λ
9.	Gamma (α,β)	$3+\frac{1}{\alpha}$	for any β
10.	Binomial (n,p)	$3-\frac{6}{n}+\frac{1}{n^2p(1-p)}$	
11.	Poisson (μ)	$3+\frac{1}{\mu}$	

Appendix A.3: Effect of the Dependence Approximation

This appendix supplements Section 2.3, pp. 40-41. We characterize the Dependence Approximation by e :

(b.1) $$v_a = v_x + e$$

Taking the mean:

(b.2) $$\bar{v}_a = \bar{v}_x + \bar{e}$$

Squaring (b.1) and taking the mean:

(b.3) $$\overline{v_a^2} = \overline{v_x^2} + \overline{e^2} + 2\overline{v_x e}$$

Squaring (b.2) :

(b.4) $$\bar{v}_a^2 = \bar{v}_x^2 + \bar{e}^2 + 2\bar{v}_x\bar{e}$$

Subtracting (b.4) from (b.3) :

(b.5) $$\check{v}_a = \check{v}_x + \check{e} + 2\mathrm{cov}(v_x, e)$$

So, from (b.2) and (b.5) , we have:

$$(\bar{v}_a - \bar{v}_x) = \bar{e}$$

$$(\check{v}_a - \check{v}_x) = \check{e} + 2\mathrm{cov}(v_x, e)$$

These equations yield the correction terms for the effect of the Dependence Approximation and are shown as Equations (2.14) and (2.15).

Appendix A.4: Calculation of $\bar{e}$, $\check{e}$ and cov(v,e)

This appendix supplements Section 2.3, p. 42. We can calculate $\bar{e}$, $\check{e}$ and $\mathrm{cov}(v_x,e)$ from the assessed conditional mean and variance of e , $\langle e|d,s,\delta\rangle$ and $^{v}\langle e|d,s,\delta\rangle$ as follows:

$$\bar{e} = \langle e|d,\delta\rangle$$

(c.1)
$$= \int_s \langle e|d,s,\delta\rangle\{s|\delta\}$$

$$\overline{e^2} = \langle e^2|d,\delta\rangle$$

$$= \int_s \langle e^2|d,s,\delta\rangle\{s|\delta\}$$

$$= \int_s \left({}^{v}\langle e|d,s,\delta\rangle + \langle e|d,s,\delta\rangle^2\right)\{s|\delta\}$$

$$\check{e} = {}^{v}\langle e|d,\delta\rangle$$

$$= \overline{e^2} - \bar{e}^2$$

(c.2)
$$= \int_s \left({}^{v}\langle e|d,s,\delta\rangle + \langle e|d,s,\delta\rangle^2\right)\{s|\delta\} - \bar{e}^2$$

$$\overline{v_x e} = \langle v_x e|d,\delta\rangle$$

$$= \int_s \int_e g(d,s)\ e\ \{s,e|d,\delta\}$$

$$= \int_s g(d,s)\{s|d,\delta\} \int_e e\{e|d,s,\delta\}$$

$$= \int_s g(d,s)\ \langle e|d,s,\delta\rangle\{s|\delta\}$$

$$\mathrm{cov}(v_x, e) = \overline{v_x e} - \bar{v}_x \bar{e}$$

$$(c.3) \qquad = \int_s g(d,s) \langle e | d,s,\delta \rangle \{s | \delta\} - \bar{v}_x \bar{e}$$

Equations (c.1) , (c.2) and (c.3) yield the required calculations and are shown as Equations (2.16) through (2.18).

Appendix A.5: Special Case of the Dependence Approximation

This appendix supplements Section 2.3, pp. 44-45. Suppose that we assess:

$$\langle e|d,v_x,\delta\rangle = A\, v_x$$

$$^{v}\langle e|d,v_x,\delta\rangle = B\, v_x^{\,2}$$

Then:

$$\bar{e} = \langle e|d,\delta\rangle = \int_v \langle e|d,v_x,\delta\rangle\{v_x|d,\delta\}$$

$$= \int_v A\, v_x\, \{v_x|d,\delta\}$$

(d.1) $$= A\,\bar{v}_x$$

$$\overline{e^2} = \langle e^2|d,\delta\rangle = \int_v \langle e^2|d,v_x,\delta\rangle\{v_x|d,\delta\}$$

$$= \int_v \left({}^{v}\langle e|d,v_x,\delta\rangle + \langle e|d,v_x,\delta\rangle^2\right)\{v_x|d,\delta\}$$

$$= \int_v (B\, v_x^{\,2} + A^2 v_x^{\,2})\{v_x|d,\delta\}$$

$$= (B + A^2)\,\overline{v_x^{\,2}}$$

$$\check{e} = {}^{v}\langle e|d,\delta\rangle = \overline{e^2} - \bar{e}^2$$

$$= (B + A^2)\,\overline{v_x^{\,2}} - A^2\bar{v}_x^{\,2}$$

(d.2) $$= (B + A^2)\,\check{v}_x + B\,\bar{v}_x^{\,2}$$

$$\overline{v_x e} = \int_v \int_e v_x e\{v_x, e|d,\delta\}$$

$$= \int_v v_x\{v_x|d,\delta\}\int_e e\{e|d,v_x,\delta\}$$

$$= \int_v v_x\{v_x|d,\delta\}\langle e|d,v_x,\delta\rangle$$

$$= \int_v A\ v_x^2\{v_x|d,\delta\}$$

$$= A\ \overline{v_x^2}$$

$$\text{cov}(v_x,e) = \overline{v_x e} - \bar{v}_x \bar{e}$$

$$= A\ \overline{v_x^2} - A\ \bar{v}_x^2$$

$$(d.3) \qquad = A\ \check{v}_x$$

Then, the correction terms for the effect of the Dependence Approximation are:

$$(\bar{v}_a - \bar{v}_x) = \bar{e} = A\ \bar{v}_x$$

$$(\check{v}_a - \check{v}_x) = \check{e} + 2\text{cov}(v_x,e) = (B + 2A+A^2)\ \check{v}_x + B\ \bar{v}_x^2$$

These equations are shown as Equations (2.22) and (2.23).

Appendix A.6: Multivariate Case

This appendix supplements Section 2.5, pp. 50-51. Assume that there are N state variables. We use the following matrix notation:

$\underline{s}$, $\underline{n}$ and $\underline{x}$ are N-dimensional column vectors

$\underline{w} = \nabla g(d,\underline{s})\big|_{\underline{s}=\bar{\underline{n}}}$, an N-dim. row vector

$W = \underline{w}^T\underline{w}$, an N x N matrix

$\underline{1}$ = an N-dim. column vector of all 1's

$L = \underline{1}\,\underline{1}^T$, an N x N matrix of all 1's

$$G_{ij} = \left.\frac{\delta^2 g(d,\underline{s})}{\delta s_i \delta s_j}\right|_{\underline{s}=\bar{\underline{n}}}$$

$G = \left[G_{ij}\right]$, an N x N matrix

$G{:}\underline{n} = \left[G_{ij}(n_i - \bar{n}_i)(n_j - \bar{n}_j)\right]$, an N x N matrix

$G{:}\underline{x} = \left[G_{ij}(x_i - \bar{n}_i)(x_j - \bar{n}_j)\right]$, an N x N matrix

$G{:}\mathrm{cov}(\underline{n}) = \left[G_{ij}\mathrm{cov}(n_i,n_j)\right]$, an N x N matrix

$G{:}\mathrm{cov}(\underline{x}) = \left[G_{ij}\mathrm{cov}(x_i,x_j)\right]$, an N x N matrix

$G{:}\Delta\bar{\underline{s}} = \left[G_{ij}\,\Delta\bar{s}_i\Delta\bar{s}_j\right]$, an N x N matrix

Note that, for a matrix M , $\underline{1}^T M\,\underline{1}$ is the sum of all of the elements of M.

Also note that:

$$(\underline{n} - \bar{\underline{n}})^T G (\underline{n} - \bar{\underline{n}}) = \underline{1}^T G{:}\underline{n}\ \underline{1} \quad , \text{ a scalar}$$

$$\overline{G{:}\underline{n}} = G{:}\mathrm{cov}(\underline{n})$$

$$\overline{G{:}\underline{x}} = G{:}\mathrm{cov}(\underline{x}) + G{:}\Delta\bar{\underline{s}}$$

We characterize the Distribution Approximation by:

$$\Delta\bar{\underline{s}} = \bar{\underline{x}} - \bar{\underline{n}}$$

$$\Delta\check{\underline{s}} = \check{\underline{x}} - \check{\underline{n}}$$

We expand $v_m = g(d,\underline{n})$ about $\bar{\underline{n}}$:

(e.1) $$v_m \simeq g(d,\bar{\underline{n}}) + \underline{w}(\underline{n} - \bar{\underline{n}}) + \tfrac{1}{2}\, \underline{1}^T G{:}\underline{n}\ \underline{1}$$

Taking the mean:

(e.2) $$\bar{v}_m \simeq g(d,\bar{\underline{n}}) + \tfrac{1}{2}\, \underline{1}^T G{:}\mathrm{cov}(\underline{n})\ \underline{1}$$

Squaring (e.1) and taking the mean:

$$\begin{aligned} v_m^2 \simeq\ & g(d,\bar{\underline{n}})^2 + \underline{1}^T W{:}\underline{n}\ \underline{1} + \tfrac{1}{4}\, \underline{1}^T\ G{:}\underline{n}\ L\ G{:}\underline{n}\ \underline{1} \\ & + 2g(d,\bar{\underline{n}})\ \underline{w}\ (\underline{n} - \bar{\underline{n}}) + g(d,\bar{\underline{n}})\ \underline{1}^T G{:}\underline{n}\ \underline{1} \\ & + \underline{w}\ (\underline{n} - \bar{\underline{n}})\ \underline{1}^T G{:}\underline{n}\ \underline{1} \end{aligned}$$

(e.3) $$\begin{aligned} \overline{v_m^2} \simeq\ & g(d,\bar{\underline{n}})^2 + \underline{1}^T W{:}\mathrm{cov}(\underline{n})\ \underline{1} + \tfrac{1}{4}\, \underline{1}^T\ \overline{G{:}\underline{n}\ L\ G{:}\underline{n}}\ \underline{1} \\ & + g(d,\bar{\underline{n}})\ \underline{1}^T G{:}\mathrm{cov}(\underline{n})\ \underline{1} + \overline{\underline{w}(\underline{n} - \bar{\underline{n}})\ \underline{1}^T G{:}\underline{n}\ \underline{1}} \end{aligned}$$

Squaring (e.2) :

$$\text{(e.4)} \qquad \bar{v}_m^{\,2} \approx g(d,\bar{\underline{n}})^2 + \tfrac{1}{4}\,\underline{1}^T\, G{:}\mathrm{cov}(\underline{n})\; L\; G{:}\mathrm{cov}(\underline{n})\;\underline{1}$$

$$+ \; g(d,\bar{\underline{n}})\;\underline{1}^T\, G{:}\mathrm{cov}(\underline{n})\;\underline{1}$$

Subtracting (e.4) from (e.3) :

$$\text{(e.5)} \qquad \check{v}_m \approx \underline{1}^T\, W{:}\mathrm{cov}(\underline{n})\;\underline{1}$$

$$+ \; \tfrac{1}{4}\,\underline{1}^T \left[\overline{G{:}\underline{n}\; L\; G{:}\underline{n}} - G{:}\mathrm{cov}(\underline{n})\; L\; G{:}\mathrm{cov}(\underline{n})\right]\underline{1}$$

$$+ \; \overline{\underline{w}\;(\underline{n} - \bar{\underline{n}})\;\underline{1}^T\, G{:}\underline{n}\;\underline{1}}$$

Now, we expand $v_x = g(d,\underline{x})$ about $\bar{\underline{n}}$:

$$\text{(e.6)} \qquad v_x \approx g(d,\bar{\underline{n}}) + \underline{w}\;(\underline{x} - \bar{\underline{n}}) + \tfrac{1}{2}\,\underline{1}^T\, G{:}\underline{x}\;\underline{1}$$

Taking the mean:

$$\text{(e.7)} \qquad \bar{v}_x \approx g(d,\bar{\underline{n}}) + \underline{w}\Delta\bar{\underline{s}} + \tfrac{1}{2}\,\underline{1}^T \left[G{:}\mathrm{cov}(\underline{x}) + G{:}\Delta\bar{\underline{s}}\right]\underline{1}$$

Squaring (e.6) and taking the mean:

$$v_x^{\,2} \approx g(d,\bar{\underline{n}})^2 + \underline{1}^T\, W{:}\underline{x}\;\underline{1} + \tfrac{1}{4}\,\underline{1}^T\, G{:}\underline{x}\; L\; G{:}\underline{x}\;\underline{1}$$

$$+ \; 2g(d,\bar{\underline{n}})\;\underline{w}\;(\underline{x} - \bar{\underline{n}}) + g(d,\bar{\underline{n}})\;\underline{1}^T\, G{:}\underline{x}\;\underline{1}$$

$$+ \; \underline{w}\;(\underline{x} - \bar{\underline{n}})\;\underline{1}^T\, G{:}\underline{x}\;\underline{1}$$

(e.8) $$\overline{v_x^{\,2}} \simeq g(d,\bar{\underline{n}})^2 + \underline{1}^T \left[W\!:\!cov(\underline{x}) + W\!:\!\Delta\bar{\underline{s}} \right] \underline{1}$$

$$+ \tfrac{1}{4}\, \underline{1}^T\, \overline{G\!:\!\underline{x}\; L\; G\!:\!\underline{x}}\; \underline{1}$$

$$+ 2g(d,\bar{\underline{n}})\, \underline{w}\Delta\bar{\underline{s}} + g(d,\bar{\underline{n}})\, \underline{1}^T \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] \underline{1}$$

$$+ \underline{w}\, \overline{(\underline{x} - \bar{\underline{n}})\, \underline{1}^T \;\; G\!:\!\underline{x}}\; \underline{1}$$

Squaring (e.7) :

(e.9) $$\bar{v}_x^{\,2} \simeq g(d,\bar{\underline{n}})^2 + \underline{1}^T\, W\!:\!\Delta\bar{\underline{s}}\, \underline{1}$$

$$+ \tfrac{1}{4}\, \underline{1}^T \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] L \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] \underline{1}$$

$$+ 2g(d,\bar{\underline{n}})\, \underline{w}\Delta\bar{\underline{s}} + g(d,\bar{\underline{n}})\, \underline{1}^T \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] \underline{1}$$

$$+ \underline{w}\Delta\bar{\underline{s}}\, \underline{1}^T \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] \underline{1}$$

Subtracting (e.9) from (e.8) :

(e.10) $$\check{v}_x \simeq \underline{1}^T\; W\!:\!cov(\underline{x})\; \underline{1} + \tfrac{1}{4}\, \underline{1}^T\; \overline{G\!:\!\underline{x}\; L\; G\!:\!\underline{x}}\; \underline{1}$$

$$- \tfrac{1}{4}\, \underline{1}^T \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] L \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] \underline{1}$$

$$+ \underline{w} \left(\overline{(\underline{x} - \bar{\underline{n}})\, \underline{1}^T \;\; G\!:\!\underline{x}}\; \underline{1} - \Delta\bar{\underline{s}}\, \underline{1}^T \left[G\!:\!cov(\underline{x}) + G\!:\!\Delta\bar{\underline{s}} \right] \underline{1} \right)$$

Subtracting (e.2) from (e.7) and (e.5) from (e.10) , we obtain the estimated correction terms for the effect of the Distribution Approximation:

(e.11) $$(\bar{v}_x - \bar{v}_m) \simeq \underline{w}\Delta\bar{\underline{s}} + \tfrac{1}{2}\, \underline{1}^T \left[G\!:\!cov(\underline{x}) - G\!:\!cov(\underline{n}) + G\!:\!\Delta\bar{\underline{s}} \right] \underline{1}$$

$$(\check{v}_x - \check{v}_m) \simeq \underline{1}^T \left[W{:}\mathrm{cov}(\underline{x}) - W{:}\mathrm{cov}(\underline{n}) \right] \underline{1}$$

$$+ \tfrac{1}{2}\, \underline{1}^T \Big(\overline{G{:}\underline{x}\; L\; G{:}\underline{x}} - \overline{G{:}\underline{n}\; L\; G{:}\underline{n}}$$

$$- \left[G{:}\mathrm{cov}(\underline{x}) + G{:}\Delta\underline{\bar{s}} \right] L \left[G{:}\mathrm{cov}(\underline{x}) + G{:}\Delta\underline{\bar{s}} \right] \qquad \text{(e.12)}$$

$$+ G{:}\mathrm{cov}(\underline{n})\; L\; G{:}\mathrm{cov}(\underline{n}) \Big) \underline{1}$$

$$+ \underline{w} \Big(\overline{(\underline{x} - \underline{\bar{n}})\, \underline{1}^T \;\; G{:}\underline{x}\, \underline{1}} - \overline{(\underline{n} - \underline{\bar{n}})\, \underline{1}^T \;\; G{:}\underline{n}\, \underline{1}}$$

$$- \Delta\underline{\bar{s}}\; \underline{1}^T \left[G{:}\mathrm{cov}(\underline{x}) + G{:}\Delta\underline{\bar{s}} \right] \underline{1} \Big)$$

These equations are not useful because they require more information than we have about the two distributions on $\underline{s}$. However, when the state variables are mutually independent, these equations reduce to a more useful form. Letting g_i and g_{ij} denote the partial derivatives of g and letting $\overset{3}{s}$ and $\overset{4}{s}$ denote the third- and fourth central moments, we have for mutually independent state variables:

$$(\bar{v}_x - \bar{v}_m) \simeq \sum_i g_i(d,\underline{\bar{n}})\Delta\bar{s}_i + \tfrac{1}{2} \sum_i g_{ii}(d,\underline{\bar{n}})\Delta\check{s}_i$$

$$+ \tfrac{1}{2} \sum_i \sum_j g_{ij}(d,\underline{\bar{n}})\Delta\bar{s}_i \Delta\bar{s}_j \qquad \text{(e.13)}$$

$$
\begin{aligned}
(\check{v}_x - \check{v}_m) \simeq\ & \sum_i g_i(d,\underline{\bar{n}})^2 \Delta\check{s}_i \\
& + \tfrac{1}{4} \sum_i g_{ii}(d,\underline{\bar{n}})^2 (\overset{4}{x}_i - \check{x}_i^2 - \overset{4}{n}_i + \check{n}_i^2 + 4\overset{3}{x}_i \Delta\bar{s}_i + 4\check{x}_i \Delta\bar{s}_i^2) \\
& + \tfrac{1}{2} \sum_i \sum_{j \neq i} g_{ij}(d,\underline{\bar{n}})^2 (\check{x}_i \check{x}_j - \check{n}_i \check{n}_j + \check{x}_i \Delta\bar{s}_j^2 + \check{x}_j \Delta\bar{s}_i^2) \\
& + \sum_i \sum_j \sum_{k \neq j} g_{ij}(d,\underline{\bar{n}}) g_{ik}(d,\underline{\bar{n}}) (\tfrac{1}{2}\overset{3}{x}_i \Delta\bar{s}_j + \check{x}_i \Delta\bar{s}_j \Delta\bar{s}_k) \\
& + \sum_i g_i(d,\underline{\bar{n}}) g_{ii}(d,\underline{\bar{n}}) (\overset{3}{x}_i - \overset{3}{n}_i + 2\check{x}_i \Delta\bar{s}_i) \\
& + \sum_i \sum_{j \neq i} g_i(d,\underline{\bar{n}}) g_{ij}(d,\underline{\bar{n}})\ 2\check{x}_i \Delta\bar{s}_j
\end{aligned}
\qquad \text{(e.14)}
$$

To eliminate the third- and fourth-order terms, we apply the simplifying assumptions (see Appendix A.2) :

$$
\overset{3}{s} = 0
$$

$$
\overset{4}{s} = 3\check{s}^2
$$

Then (e.14) becomes:

$$
\begin{aligned}
(\check{v}_x - \check{v}_m) \simeq\ & \sum_i g_i(d,\underline{\bar{n}})^2 \Delta\check{s}_i \\
& + \tfrac{1}{2} \sum_i \sum_j g_{ij}(d,\underline{\bar{n}})^2 (\check{x}_i \check{x}_j - \check{n}_i \check{n}_j + \check{x}_i \Delta\bar{s}_j^2 + \check{x}_j \Delta\bar{s}_i^2) \\
& + \sum_i \sum_j \sum_{k \neq j} g_{ij}(d,\underline{\bar{n}}) g_{ik}(d,\underline{\bar{n}})\ \check{x}_i \Delta\bar{s}_j \Delta\bar{s}_k \\
& + 2\sum_i \sum_j g_i(d,\underline{\bar{n}}) g_{ij}(d,\underline{\bar{n}})\ \check{x}_i \Delta\bar{s}_j
\end{aligned}
\qquad \text{(e.15)}
$$

Equations (e.13) and (e.15) yield the estimated correction terms for the effect of the Distribution Approximation for independent state variables and are shown as Equations (2.27) and (2.28).

REFERENCES

1. Arons, Arnold B. and Alfred M. Bork, eds., Science and Ideas, Prentice-Hall, Inc., 1964.

2. Ashby, W. Ross, "Analysis of the System To Be Modeled," The Process of Model-Building in the Behavioral Sciences, Ralph M. Stogdill, ed., Ohio State University Press, 1970, pp. 94-114.

3. Carnap, Rudolf, "Statistical and Inductive Probability" and "Inductive Logic and Science," The Galois Institute of Mathematics and Art, 1955.

4. de Finetti, Bruno, "Does it Make Sense to Speak of 'Good Probability Appraisers'?," The Scientist Speculates, I. J. Good, ed., Heinemann, 1962, pp. 357-364.

5. de Finetti, Bruno, "Probability: Interpretations," International Encyclopedia of the Social Sciences, D. L. Sills, ed., MacMillan and Free Press, 1968, vol. 12, pp. 496-505.

6. Good, Irving John, The Estimation of Probabilities, M.I.T. Press, 1965.

7. Howard, Ronald A., "Decision Analysis: Applied Decision Theory," Proceedings of the Fourth International Conference on Operational Research, Wiley-Interscience, 1966, pp. 55-71.

8. Howard, Ronald A., "The Foundations of Decision Analysis," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 3, September 1968, pp. 211-219.

9. Howard, Ronald A., "Decision Analysis: Perspectives on Inference, Decision and Experimentation," IEEE Special Issue, Detection Theory and Applications, May 1970, pp. 622-643.

10. Howard, Ronald A., "Decision Analysis in Systems Engineering," Lecture, 1971 California Institute of Technology Systems Seminar Series, April 1971.

11. Howard, Ronald A., "Proximal Decision Analysis," Management Science, Theory, Vol. 17, No. 9, May 1971, pp. 507-541.

12. Howard, Ronald A., "The Science of Decision Making," unpublished.

13. Howard, Ronald A., "Risk Preference," unpublished.

14. Hunt, Earl B., "The Evaluation of Somewhat Parallel Models," Mathematical Explorations in Behavioral Science, F. Massarik and

P. Ratdosh, eds., Irwin-Dorsey Series in Behavioral Science, 1965, pp. 37-55.

15. Kemeny, John G., A Philosopher Looks at Science, D. Van Nostrand Company, Inc., 1959.

16. Kyburg, Henry E., Jr. and Howard E. Smokler, eds., Studies in Subjective Probability, John Wiley and Sons, 1964.

17. Lempers, Frederik Burchard, "Posterior Probabilities of Alternate Linear Models," Universitaire Pers Rotterdam, 1971.

18. Matheson, James E., "The Economic Value of Analysis and Computation," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 3, September 1968, pp. 325-332.

19. Matheson, James E., "Decision Analysis Practice: Examples and Insights," Proceedings of the Fifth International Conference in Operational Research, Venice, 1969, Tavistock Publications. Also available as SRI Publication DAG-16.

20. Morris, Peter Alan, Bayesian Expert Resolution, Ph.D. dissertation, Department of Engineering-Economic Systems, Stanford University, May 1971.

21. Murray, George R., Jr., "Modeling and the Use of Subjective Knowledge in Engineering Design and Experiment Selection," unpublished.

22. Pollock, Stephen M., "The Modeling Studio," unpublished.

23. Raiffa, Howard, Decision Analysis: Introductory Lectures on Choice Under Uncertainty, Addison-Wesley, 1968.

24. Raiffa, Howard, "Assessments of Probabilities," unpublished.

25. Raiffa, Howard and Marc Alpert, "A Progress Report on the Training of Probability Assessors," unpublished.

26. Smallwood, Richard D., "A Decision Analysis of Model Selection," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 3, September 1968, pp. 333-342.

27. Spetzler, Carl S. and Ramon M. Zamora, "Decision Analysis of a Facilities Investment and Expansion Problem," SRI Publication DAG-46.

28. Spetzler, Carl S. and Carl-Axel S. Staël von Holstein, "Probability Encoding in Decision Analysis," presented at ORSA-TIMS-AIEE 1972 Joint National Meeting, November 1972.

29. Tversky, Amos and Daniel Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," SRI Publication.

30. Winkler, Robert L. , "The Assessment of Prior Distributions in Bayesian Analysis," American Statistical Association Journal, September 1967, pp. 776-800.