Statistical Decision Theory

Benny R. Copeland

This new approach to problem solving can lead to better decisions if properly applied. However, the technique can only complement, not replace, the executive's own knowledge and experience.

STATISTICAL DECISION THEORY

by Benny R. Copeland
North Texas State University

Business theoreticians—and in a few cases hard-headed businessmen—until a few years ago employed scientific approaches only to very specific problems, such as inventory level. In recent years, however, a general decision making algorithm (methodology) has been developed which has quite wide application to business problems. As a matter of fact, it is the generality or universal applicability of this algorithm that makes it so very significant.

Basically, the very essence of the management process is decision making. Thus a decision making algorithm—a rule which can be expressed in mathematical terms—in effect is a description of this aspect of the management process. As such, it should serve to complete the philosophical theory necessary for truly scientific management.

The title of this article emphasizes *statistical* decision theory. This modifier was added in recognition of the fact that quantitative methods are today essential to the stated expression of business methods and policies. Certainly this does not make the algorithm less general. Nothing within our universe is more general than mathematics. The purpose of this article is to examine the essence of statistical decision theory and to indicate its application by means of an illustrative example.

The methodology

An appropriate place to begin our investigation of decision theory would seem to be with a definition of the term. By "decision theory" we shall mean an algorithm which results in the selection of the proper action to be taken in a decision situation from among many alternative actions. By implication we have also defined what is meant by "decision making," i.e., selecting the best alternative action.

Defining a term properly is not easily accomplished. Let us examine the algorithm in detail:

The Decision Making Algorithm

A. Define the problem.
B. Develop the appropriate decision criteria.
C. Determine the environmental situation.
D. Describe all possible actions.
E. Develop the decision model.
F. Solve the model.
G. Make the decision.

Each step of this methodology will be explored in detail below.

A decision is characterized by a goal, the availability of several possible actions, and the environment of the outcomes of the actions (certainty, uncertainty, risk, conflict, ignorance).

Define the problem

A decision problem is characterized by:

A. The desire to attain a certain goal
B. The availability of several actions which can be taken, some of which will not be as effective as others
C. The particular type of environment which exists with respect to the action outcomes (certainty, uncertainty, risk, conflict, ignorance)

It is of prime importance that the decision maker analyze his problem in terms of each of the above characteristics. By gaining a better understanding of his problem he simplifies its solution.

For illustrative purposes the following discussion of decision theory will be built around a highly simplified inventory problem. Especially note that the problem statement is directed toward the attainment of a particular goal.

> During summer vacation a high school student sells cut roses by the dozen at a roadside stand. The roses cost the student $1 a dozen and sell for $3 a dozen. Because the stand has no refrigeration facilities the roses must be purchased fresh each day from the wholesaler. *What is the most economic order quantity for the student?* Upon inquiry he supplies the following data on past demand:

Demand (*in dozens*)	Days (*of demand*)
0	2
1	3
2	4
3	1

Analyzing this problem in terms of the characteristics set forth above we find:

A. The goal—how many dozen roses should the student purchase each morning? (The decision maker will generally find it useful to state the problem as a question.)
B. The set of actions which the student can take corresponds to the various inventory levels he should stock: 0, 1, 2, or 3 dozen roses.
C. Comparison of the problem situation with the various classes of environment indicates that the problem involves decision making under risk (because past experience provides a useful probability distribution of outcomes for each action.)

Develop criteria

A decision criterion is an indicator or index that would serve as an appropriate means of measuring attainment of the goal. For the problem at hand we might ask ourselves, "What measure denotes the

BENNY R. COPELAND, Ph.D., CPA, is associate professor of business administration at North Texas State University. He is a member of the American Institute of CPAs, the American Accounting Association, and American Statistical Association. Dr. Copeland's articles have appeared in this publication as well as **The Accounting Review**, **Business Studies**, and **Financial Executive**.

proper inventory level?" The possible criteria might include the following:

A. We sell all inventory every day.
B. We never have to turn down a sale for lack of inventory.
C. We maximize short-run profit.
D. We maximize long-run profit.

Selection of a decision criterion involves the making of a subjective or "value" judgment. By nature, value judgments are of a short-term or *ad hoc* nature, and therefore they tend to vary with the nature of the problem. Some of the more commonly applied criteria for business decision problems which have been explored in the literature of decision theory are:

A. Maximum absolute gain
B. Maximum expected gain
C. Minimum expected loss
D. Minimum absolute loss
E. Maximum expected net value

"Maximum absolute gain" is the criterion of the complete optimist (and of the complete gambler), that individual who must always "go for broke." This criterion considers only the magnitude of the profits of each action and would result in selecting that action with the largest absolute profit—regardless of the probability of attaining that profit. (Maximize P_j.)

"Maximum expected gain" as a criterion considers not only the absolute size of the potential profit for each action but also the related probability of attainment. Thus a $10,000 profit with a 50 per cent probability of attainment would be exactly as desirable as a $20,000 profit with a 25 per cent probability of attainment. (Maximize EP_j.)

"Minimum expected loss," as might be anticipated from the previous discussion, involves selecting the action with the smallest EL_j* or expected loss. This criterion involves the consideration of possible losses and their related probabilities of occurrence. On the other hand, the following criterion, "minimum absolute loss," considers only the absolute magnitude of the possible losses and selects the smallest. This criterion will minimize losses if the worst possible event occurs. Because the measurement is absolute it gives no consideration to the probabilities of occurrence of the losses; the smallest L_j is chosen because it represents the smallest loss, and no attempt is made to calculate the probability of its occurrence.

"Maximum expected net value" is defined as the expected value of the profits minus the expected value of the losses. This criterion considers all possible elements of the problem and is thus probably the most comprehensive measure. Suppose that a particular action has 60 per cent probability of producing a profit of $12,000 and a 40 per cent probability of producing a loss of $5,000; the expected net value of this action is computed as follows:

$$\begin{aligned} \text{E.N.V.} &= EP_j - EL_j \\ &= .6(\$12{,}000) - .4(\$5{,}000) \\ &= \$7{,}200 - \$2{,}000 \\ &= \$5{,}200. \end{aligned}$$

Throughout the present discussion of decision criteria we have considered their measurement only in terms of dollars. Certain value judgments, however, cannot readily be stated in dollars, such as criterion B in our example which stated, with respect to the illustrative problem, that we desired to carry a level of inventory such that we would never have to turn down a sale for lack of inventory. For non-monetary criteria such as this, or for measuring those instances when the firm's utility function for money is not linear, it becomes necessary to set criteria in "utiles" rather than dollars (i.e., to weight dollars before taking them into consideration). This area is known as "utility theory" and falls outside the scope of this paper. The bibliography will refer the reader to selected references if he wishes to pursue the

Selection of a decision criterion involves the making of a subjective or "value" judgment. By nature, value judgments are of a short-term or ad hoc nature and tend to vary with the nature of the problem.

* The symbol E is read "expected value of," thus EP_j is read "expected value of the various profits." The subscript $_j$ is read for the level of inventory. See Exhibit A on page 50 for a discussion of this concept.

The introduction of intuitive or "subjective" probabilities (as opposed to "objective" or empirical probabilities) into the decision making algorithm has precipitated long and heated argument among statisticians.

study of this topic. (See the listing on page 51.)

Determine the environment

If we define decision making as the selection of one alternative from among several, we may then identify the specific decision making situations typical for the business executive:

A. Certainty—The set of all alternative actions is known, and the outcome of each action is known with certainty.

B. Risk—The set of all alternative actions is known, but the outcome of each action can be stored only in terms of a probability distribution.

C. Uncertainty—The set of all alternative actions is known, but the outcome of each action is uncertain.

D. Conflict—The set of all alternative actions is known, but the outcome of each is dependent upon the reaction of a knowledgeable opponent.

E. Ignorance—The set of all alternative actions is unknown.

The above list of decision situations is in the order of desirability. Ideally we would like to operate always under the condition of certainty. Yet very rarely is this condition faced by the executive. When the situation is encountered by the executive it is usually in the area of production management. One example of decision making under certainty is the problem of setting product mix so as to maximize the allocation of capital goods. Linear programing is a statistical technique applicable to problems of this class.

Decision making under risk is characterized by the availability of historical probability distributions for the outcomes of the various actions. For example, if we take a particular action with respect to adjusting a machine we know that the outcome will be 5 per cent defectives and 95 per cent non-defectives. For this information to be available it is necessary, of course, that the decision situation be of a repetitive nature. Many types of business decision situations are of this nature.

When the situation is one of decision making under uncertainty the first determining characteristic is that historical probabilities are not available to the decision maker. This may be because the process is of a non-repetitive nature, or, it may simply be because no data have been collected. The second characteristic is that the decision maker has, or can obtain, an intuitive concept of the situational probabilities. Unless this can be done the situation becomes one of "ignorance," and a logical solution becomes impossible to approach.

The admission of intuitive or "subjective" probabilities (as opposed to "objective" or empirical probabilities) into the decision making algorithm has instigated a long and heated argument among statisticians. Those in favor of admitting the subjective probabilities are referred to as "Bayesians," named after Thomas Bayes (1702-1761), a Presbyterian minister at Tunbridge Wells in England. An essay of Bayes, published posthumously in 1763, offered a theorem for finding the inverse probability of an event. Although this theorem, as developed, was directed toward classical probability theory, it has since been adopted by "Bayesians" for purposes of merging subjective probabilities with subsequently obtained sampling information. In its classical form Bayes' Theorem is:

$$P(X/A) = [P(X) \cdot P(A/X)]$$

divided by

$$[P(X) \cdot P(A/X) + P(\tilde{X}) \cdot P(A/\tilde{X})]$$

$$= [P(X) \cdot P(A/X)]$$

divided by

$$P(A \quad X) + P(A \quad \tilde{X})$$

$$= [P(X) \cdot P(A/X)]$$

divided by $P(A)$

The classical function of this theorem was to find $P(X/A)$ given $P(A/X)$, thus we sometimes find this theorem referred to as the "theorem of inverse probability." As adopted by Bayesians, the theorem becomes:

$$\text{Posterior Probabilities} = \frac{[\text{Subjective Probabilities times Sample Problems}]}{\sum [\text{Subjective Probabilities times sample Problems}]}$$

The purpose of Bayes' Theorem when applied by the Bayesians is to provide a formal algorithm for adjusting one's opinion in light of additional data. It has been charged, and perhaps with some justification, that the revision (Bayesian) methodology has little or nothing to do with Bayes' Theorem, that in effect the process in actuality produces a weighted arithmetic mean of the form:

$$P_{j_1} = \frac{[P_j \cdot \text{Specific Weight}]}{\sum_{j=1}^{N} [P_j \cdot \text{Specific Weight}]}$$

Regardless of the methodology, the "non-Bayesian" or "classical" statistician recoils with horror from admitting subjective probabilities into a statistical process. His position is: "If historical probabilities are not available, there is nothing for the statistician to work with; he can (and should) do nothing."

This article does not intend to take a position either way on the Bayesian question. An attempt has been made to identify the area of disagreement—nothing more. It is only fair to note, however, that if subjective probabilities are not allowed to be considered then it becomes impossible for the executive to make rational decisions under conditions of uncertainty. And, this process also demands that experienced business executives have developed a general "feeling" for their job.

For some reason "decision theory" seems, by usage, to have become somewhat synonymous with "decision making under uncertainty." This is not the meaning of "decision theory" as used within this article, and this trend should perhaps be resisted. Otherwise we will have no simple term to refer to all of the decision situations described above, and we shall have to go to the trouble of inventing a new term to refer to the generic process.

Describe possible actions

The importance of step C., "describe all possible actions which can be taken," is patently obvious—it alerts the decision maker to his entire set of alternatives. The necessity for having this list as complete as possible cannot be stressed too much, for if the most appropriate action is excluded from this list whatever decision is reached may not be the most efficient solution.

The difficulty of describing the universe of actions will of necessity vary with the problem. In the illustrative problem under present consideration the actions open to the decision maker are particularly simple to develop:

Action	*Order Quantity*
A	0
B	1
C	2
D	3

The reader must be cautioned that this simplicity is unusual. Developing the universal set of actions for most real-world problems is generally of a more complex nature.

Describe the outcomes

Given the selected decision criterion and the universe of available actions, the next step of the decision maker is to evaluate alternative actions. This evaluation is accomplished by means of a mathematical model designed to express the relationship between the environment and the actions in terms of the preselected decision criterion. Each problem situation is unique and requires development of the appropriate model. Thus, no generalization can be made as to appropriate models.

The illustrative example concerns the determination of the most efficient economic order quantity as measured by maximum expected net value. The appropriate decision model for this situation is:

Maximize $E(P_j - L_j)$

Where: P_j = the gross profit of jth level of inventory, i.e., the sales prices of the roses sold less the cost of all roses purchased.

L_j = the opportunity losses associated with each jth stock level.

This model may not be completely clear at the moment. We shall return to its meaning at length a bit later. At the present the important factor to note is that the model is developed by the decision maker from his knowledge of the problem area. Unfortunately, statistics cannot remove the need for the decision maker to know the problem area thoroughly. At best the addition of statistical techniques serves to make consideration of the problem more precise.

It is well to note at this point that there may be more than just one satisfactory decision model. The model is a conceptualization of the relationship between the actions and the environment. It is possible to examine this relationship from various points of view. Development and selection of the model are an expression of a value judgment as much as selection of the decision criterion. Various models will result in different "answers," i.e., the selection of different actions. Each "answer" is "correct" when viewed in terms of the model used. Again the point should be made that the decision maker must have a thorough knowledge of the problem area in order to apply statistical decision theory.

Solving the model

The decision model previously determined was:

Maximize: $E(P_j - L_j)$

The P subscript, j, refers to the various profits in the model—one for each inventory level. Thus P_1 is the profit from stocking one unit, P_2 is the profit from stocking two units, etc. L_j is to be read in a similar fashion. The E is read "expected value" as previously explained. The computations indicated by the model will be made clear as the solution progresses.

The first step in the solution is to compute P_j or gross profit for each inventory level. Because this gross profit is dependent upon the inventory level, it is referred to as "conditional" gross profit. When the P_j values are inserted into a table or "matrix" of the form below they are referred to as a "pay-off" matrix.

Pay-Off Matrix Showing Conditional Gross Profit (P_j) (in dollars)

Demand	Inventory level			
(0)	0	1	2	3
0	0	−1	−2	−3
0	0	+2	+1	0
2	0	+2	+4	+3
3	0	+2	+4	+6

Values in the pay-off matrix were computed in this way:

Situation: Demand, two dozen—
Stock, three dozen

Gross revenue (2 x \$3)	\$6.00
Cost of sales (3 x \$1)	3.00
Gross profit	\$3.00

The model requires that we also consider opportunity losses, however. Before we can begin to compute these values we must first define what we mean by the term "opportunity loss."

By "opportunity loss" we shall mean "a foregone benefit."

Following this definition we shall measure opportunity loss by "foregone profit," i.e., the \$2 gross profit lost each time a demanded product unit is not on hand. Applying this measurement the following matrix is obtained:

Matrix Showing Conditional Opportunity Loss (L_j) (in dollars)

Demand	Inventory level			
(D)	0	1	2	3
0	0	0	0	0
1	−2	0	0	0
2	−4	−2	0	0
3	−6	−4	−2	0

The model is based upon the concept ($P_j - L_j$). Thus far we have the conditional values computed but have not yet merged them to form the conditional net value. Simple subtraction is applied to derive the following matrix showing the conditional value ($P_j - L_j$).

Matrix Showing the Conditional Value ($P_j - L_j$) (in dollars)

Demand	Inventory level			
(D)	0	1	2	3
0	0	−1	−2	−3
1	−2	+2	+1	0
2	−4	0	+4	+3
3	−6	−2	+2	+6

The final step in solving the model is to convert the conditional (absolute) values into expected values by multiplying through by the respective probabilities of each level of demand. It was previously

EXHIBIT A

Mathematical expectation can most easily be described by means of a simple example. Suppose we play this game: We agree to flip a perfectly balanced coin. If "heads" comes up, we receive \$1.00, but if "tails" comes up, we lose \$2.00. The expected value of this game is found by multiplying the value of success times the probability of success and subtracting the product obtained by multiplying the value of failure times the probability of failure. Mathematically:

$$P_rS(S) \quad - \quad P_rF(F) \quad = \text{Expected Value}$$

$$(.50)\ (\$1.00) \quad - \quad (.50\ (\$2.00) \quad =$$

$$\$.50 \quad - \quad \$1.00 \quad = -\$.50$$

This value, —\$.50, is the average benefit we should expect to obtain if we repeated this game many times. Conceptually, mathematical expectation is a postulate which states the philosophical assumption that absolute values and probabilities may be joined through multiplication to produce meaningful values. Perhaps it might be useful, in addition, to think of expected value as a "weighted average," where the weights are probabilities and the denominator is the implicit sum of the probabilities, unity or one.

determined that the historical probabilities would be used. These were:

Demand (D)	Probability (P)
0	.20
1	.30
2	.40
3	.10
	1.00

If we multiply the $(P_j - L_j)$ matrix by the probability distribution we obtain the final solution matrix of the form $E(P_j - L_j)$, shown in Exhibit B on this page.

This matrix shows that the expected value $(P_j - L_j)$, as defined, is maximized by stocking two dozen roses each day. However, before we advise the student to act accordingly, let us note the appropriate characteristics of this solution:

A. The problem situation was characterized as a "risk" environment, and empirical probabilities were used. Just how good were these figures? Do they still apply?

B. The criterion was a value judgment in view of society's demands, long-run profit needs, short-run cash needs, firm objectives, etc.

C. The decision model was a conceptualization of the appropriate relationship between environment and actions. Was it valid?

D. For the decision to be maximally effective the list of alternative actions must contain the "best" action. Did it?

If the decision maker is satisfied regarding all the above characteristics, then the decision theory algorithm gave a "best" answer.

Solution Matrix Showing the Expected Value of $(P_j - L_j)$ (in dollars)

Demand (D)	Inventory level 0	1	2	3
0	0	−.20	− .40	− .60
1	− .60	0	+ .30	0
2	−1.60	+.60	+1.60	+1.20
3	− .60	−.20	+ .20	+ .60
$E(P_1 - L_1)$	−2.80			
$E(P_2 - L_2)$		+.20		
$E(P_3 - L_3)$			+1.70	
$E(P_4 - L_4)$				+1.20

EXHIBIT B

Conclusion

An attempt was made to define the generic decision theory process and to note its relevance to the solution of business problems. The major points in this thesis may be summarized as follows:

A. Decision theory is a broad discipline which includes decision making under certainty, risk, uncertainty, conflict, and ignorance.

B. The decision theory algorithm leads to a best solution only if many complex value judgments have been properly made. These value judgments must be made in light of a thorough knowledge of the problem area. Statistical decision theory complements experience and knowledge of business; it does not and cannot replace it.

SELECTED REFERENCES

Books

Aigner, Dennis J., *Principles of Statistical Decision Making,* The Macmillan Company, New York, 1968.

Bierman Jr., Harold and others, *Quantitative Analysis for Business Decisions* (Revised Edition), Richard D. Irwin, Inc., Homewood, Illinois, 1965.

Fellner, William, *Probability and Profit,* Richard D. Irwin, Inc., Homewood, Illinois, 1965.

Hadley, G., *Introduction to Probability and Statistical Decision Theory,* Holden-Day, Inc., San Francisco, 1967.

Luce, R. Duncan and Howard Raiffa, *Games and Decisions,* John Wiley & Sons, Inc., New York, 1957.

Schlaifer, Robert, *Introduction to Statistics for Business Decisions,* McGraw-Hill Book Company, Inc., New York, 1961.

Schlaifer, Robert, *Probability and Statistics for Business Decisions,* McGraw-Hill Book Company, Inc., New York, 1959.

Weiss, Lionel, *Statistical Decision Theory,* McGraw-Hill Book Company, Inc., New York, 1961.

Articles

Anscombe, F. J., "Bayesian Statistics," *The American Statistician,* February, 1961, pp. 21 - 24.

Freund, John E. and Irwin Miller, "In Defense of Frequencies," *The American Statistician,* April, 1955, pp. 15 - 17 and 27.

Hartley, H. O., "In Dr. Bayes' Consulting Room," *The American Statistician,* February, 1963, pp. 22 - 24.