

APPLICATION OF STATISTICAL DECISION
THEORY TO LEGAL DECISIONS

by

ROBERT HARRY MORSE

S.B., Massachusetts Institute of Technology
(1963)

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 1964

Signature of Author.....
Department of Electrical Engineering, August 21, 1964

Certified by.....
Thesis Supervisor

Accepted by.....
Chairman, Departmental Committee on Graduate Students

APPLICATION OF STATISTICAL DECISION
THEORY TO LEGAL DECISIONS

by

ROBERT HARRY MORSE

Submitted to the Department of Electrical Engineering on August 24, 1964 in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

ABSTRACT

In this thesis the author discusses generally the theory of making decisions using statistical methods, i.e., statistical decision theory. He then proceeds to build a model in which he applies this statistical decision theory to the specific application of making legal decisions.

After the model is discussed and illustrated with a theoretical legal decision the author, through exhaustive interviews with members of the legal profession, presents an actual law case which he analyzes in accordance with his theoretical model.

In a chapter entitled "The Problems" the author explains the difficulties in applying a scientific theory to analyze a real world situation. He also explains his approach in overcoming the inherent difficulties present when one tries to obtain scientific data to be used in a scientific model from an unscientific source (the attorney).

The basic gain to the attorney in using the techniques of this thesis is that he can evaluate the usefulness of any method he may wish to employ to increase his knowledge of the "unknown factors" which are hindering his decision making ability. By application of a Bayesian analysis of the statistical data based on the past experiences of the attorney, the statistical effect of any above-mentioned method is calculated.

This thesis is applied only to very simple decisions here, but its extension to more complex legal decisions with the use of computers is also discussed.

Thesis Supervisor: Ronald A. Howard
Title: Associate Professor of Electrical Engineering

ACKNOWLEDGMENTS

Mr. Stuart Macmillan,

for his tireless effort on my behalf, for his deep insight into the nature of legal decisions and for the many hours of conferences we have had, I am forever grateful.

Mr. Joseph Caulfield,

for his insight and aid in the statistical analysis of legal decisions, and for the many sessions that he afforded me, I am forever grateful.

Mrs. Margaret G. Higgins,

for her patience and fortitude in the typing and retyping of this work, and for her expert and tireless effort in my behalf, I am forever grateful.

Kenway, Jenney & Hildreth,

for this law firm's indulgence and patience, for their encouragement and interest and for their invaluable assistance, I am forever grateful.

Professor Ronald A. Howard,

for my advisor's constant vigilance and advice, for his keen interest and sympathy with the problem involved, and for his aid and assistance, I am forever grateful.

TABLE OF CONTENTS

	<u>Page</u>
Abstract	2
Acknowledgments	4
Introduction	6
Chapter I, <u>The Theory</u>	8
Chapter II, <u>The Problems</u>	17
Chapter III, <u>A theoretical Example</u>	21
Chapter IV, <u>A Practical Example</u>	35
Chapter V, Conclusion and Further Suggestions	57
Appendix	61
Bibliography	67
Figure 1	16
Figure 2 and 2a	25
Figure 3	30
Figure 4	45
Figure 5	52
Table 1 and 1a	24
Table 2	44
Table 3	49

APPLICATION OF STATISTICAL DECISION
THEORY TO LEGAL DECISIONS

INTRODUCTION

Decision making may assume all shapes and dimensions. It may be the smallest of everyday decisions that are made, usually without much thought of process or consequence, or it may be the decision of the President of the United States which is made with careful thought both to the information inputs and the consequences of the action chosen. The latter type of decision more visually points up the use of a systematic process in decision making, whereas the former may not be worth the trouble to systematize. However, before you can decide which decisions are worth systematic analysis you must have some framework in which you can compare various decisions. Such a framework is the theory of statistical decision making.

Statistical Decision Theory, based upon the work of Wald after the last World War, created an object of attention for a wide class of mathematicians, philosophers and scientists, and led to a general revamping of the then current approaches to decision problems.

Statistical decision theory is concerned with the mathematical analysis of decision making when the "state of nature" is unknown but further knowledge about the true "state

of nature" may be obtained through experimentation, possibly at a cost. The object of statistical decision theory is to choose a course of action, which may or may not attempt to improve our knowledge of the true "state of nature", which is consistent in some way with the decision maker's own personal preferences, as expressed through his utility curves and which is also consistent with what weight he assigns the possible "states of nature", as expressed by numerical probabilities.

In applying statistical decision theory, we realize that if the true "state of nature" is known with absolute certainty, then our decision making becomes trivial, and therefore, any meaningful decision process will be statistical in nature. In addition to these probabilistic considerations in decision making, one also relies on what is usually called experience. This is nothing more than using the lessons of the past to help in making present decisions. Some fields in which decisions are to be made lend more readily to using past experience or precedence. Legal decisions are of such a nature.

In this thesis we will investigate the application of statistical decision theory to making legal decisions.

CHAPTER I
THE THEORY

The obvious questions at this point are: just what constitutes statistical decision theory and how in the world can it be applied to so vague and sacrosanct a thing as a lawyer's decision making process in making legal decisions for his client? First of all, statistical decision theory is the analysis of the high art of making decisions in the face of an uncertain world in which further information about the uncertain world can be obtained. Further, it is the mathematical analysis utilizing the techniques of statistics and the concepts of probability theory.

Functionally the decision maker must choose a particular course of action to follow, such a course of action is to be chosen from amongst a group of potential courses of action that he could follow. Therefore, on the surface, it appears that the decision maker must choose the particular course of action that will be the best for him in some sense. Now, this would not be a problem if the "state of the world" were known with absolute certainty, because the decision maker, with a prescribed "goal" (e.g., making money) in mind, would simply peruse the possible courses of action and pick that course of action which would best (e.g., give him the most money) lead him to his "goal". Symbolically, we refer to a particular course of action by the symbol A_i (the i^{th}

course of action of those available) and we refer to a particular "state of nature" by the symbol θ_i (the i^{th} "state of nature" of those possible). However, the exact "state of nature" is not known with absolute certainty and therein lies the rub. The possible "states of nature" can then be considered as comprising a sample space (i.e., any collection of events related in some way) and a probability measure can be defined over this sample space. In other words, instead of saying that "the 'state of nature' is such and such", we say that "the 'state of nature' has a probability of being such and such." In symbols we refer to the probability of the state of nature being, say, θ_i is $P_r(\theta_i)$. Ideally, we would like $P_r(\theta_i)$ to be 1 for a particular state of nature and 0 for all the others. (See Appendix for some elementary probability theory to be used in this text.) This, of course, corresponds to knowing the exact "state of nature" with absolute certainty and therefore would correspond to perfect decision making with respect to an uncertain "state of nature".

Now the question arises, if it is so advisable to know what is the exact "state of nature" with as much certainty as possible, i.e., having a sample space which has one particular "state of nature" with a probability most nearly 1, and others with a probability most nearly 0, then how can we alter the situation where we are not at all sure what the exact "state of nature" is; i.e., a sample space with no one particular state of nature having a higher $P(\theta_i)$ than any of the

others. The answer is that we can change our knowledge about what is the true "state of nature" by gaining additional information through experimentation. In other words, we perform an experiment which has one of many outcomes and based on each of these outcomes, we have a new probability measure defined on our sample space (one probability measure for each possible experimental outcome). Symbolically, we refer to the possible outcomes of our experiment e , by Z_i . We refer to the probabilities of the possible states of nature before we perform the experiment, $P(\theta_i)$, as the a-priori probabilities, and we refer to the new probabilities of the possible states of nature after we experiment, $P(\theta_i/Z_i)$, as the a-posteriori probabilities. Note that the probabilities of the possible states of nature after the experiment is performed are conditional probabilities, conditioned on the outcome of the experiment chosen.

For those readers who may be familiar with the science of statistical communications and more specifically with the work of Prof. C.E. Shannon and his information theory, I would like to point out the analogy between the above process of increasing one's knowledge of the true "state of nature" through experimentation, thus changing the a-priori probabilities $P(\theta_i)$ into the a-posteriori probabilities $P(\theta_i/Z_1)$ on the "states of nature" sample space, and the process of changing the a-priori probability of a message sample space into the a-posteriori probabilities by transmitting the signal over a noisy transmission path and making a reception. In the latter case we

are increasing one's knowledge about which one of a set of signals was sent on a noisy path by making a reception; therefore, the a-posteriori probabilities of which signal was sent is also a conditioned probability, and it is conditioned on what was received.

So we can see that in the statistical communication case the message space is analogous to the "states of nature" space and which message was picked from this space is analogous to the exact "state of nature". The experiment is transmitting the message which was picked and receiving a signal. Now based upon the received signal (the outcome of the experiment), we can better say which message was picked (i.e., better say what is the exact "state of nature"). This completes our analogy.

Let us now complete the decision making framework by examining that aspect of it which differentiates it from the above analogous communication process. I am referring to the concept of utility.

Utility used in this text means the usefulness (or lack of usefulness with respect to a 0 utility level) of a combination of circumstances. It is the "payoff" for playing the game of decision making, and it also includes the costs or fees involved in playing the game.

Once the decision maker has chosen a particular experiment to perform in order that he may better judge what the exact state of nature is he now has a list of all the possible outcomes of this experiment. He then, according to his own preferences, assigns a utility (a gain or loss to him of some definite or indefinite quantity) to observing one of the possible outcomes Z_i of the chosen experiment and thereupon taking a particular course of action a_i . Of course we realize that the price of the experiment chosen could conceivably depend on the outcome Z_i .

We have heretofore considered two types of probability measures and both have been defined on what we have referred to as the "state of nature" sample space. These are the a-priori probabilities $P(\theta_i)$ and the a-posteriori probabilities $P(\theta_i/Z_i)$ which are conditioned on one of a set of possible outcomes of our chosen experiment. However, these outcomes Z_i are also events and as such they also make up a sample space (let us refer to this sample space as the "experiment-outcomes" sample space). Therefore, as in the "states of nature" sample space, we may assign a probability $P(Z_i)$, i.e., the probability of a certain outcome of the chosen experiment before the exact state of nature is known (an a-priori probability) and we may assign a probability $P(Z_i/\theta_i)$, i.e., the probability of a certain outcome of the chosen experiment given that the exact state of nature is θ_i (an a-posteriori probability conditioned on the exact state of nature).

All four of these probabilities, $P(\theta_i)$, $P(\theta/Z_i)$, $P(Z_i)$ and $P(Z_i/\theta_i)$ are related as explained in Appendix by the Bayesian relationship:

$$P(\theta_i) P(Z_i/\theta_i) = P(Z_i) P(\theta_i/Z_i) = P(Z_i, \theta_i)$$

where $P(Z_i, \theta_i)$ is the probability of the joint event that when you perform the chosen experiment that the outcome will be Z_i , and that the exact state of nature would be θ_i .

($P(Z_i, \theta_i)$ is thus a joint probability.)

Let us now consider how we assess the above probabilities. $P(\theta_i)$, which is the a-priori probability of the existing state of nature, is a very objective guess as to what is the present state of nature before you have performed an experiment to aid you; it is your best guess.

$P(Z_i)$, which is the probability that the outcome of a given experiment will be Z_i before the exact state of nature is known, may be exceedingly difficult to estimate and will have to be computed from the joint $P(\theta_i, Z_i)$, or based on extensive experience data. The probability $P(Z_i/\theta_i)$, which is the probability that the experiment will have an outcome Z_i when the state of nature is known to be θ_i when the experiment is performed, is an estimate sometimes based on intuition but also based on experience data which lists what the outcome of the experiment is when the state of nature is known to be θ_i . Then we may associate the relative frequency of the number of outcomes which are Z_i with the $P(Z_i/\theta_i)$.

$P(\theta_i/Z_i)$, which is the a-posteriori probability of the existing state of nature after you have experimented, can be calculated from the above-mentioned Bayesian relationship when you know $P(\theta_i)$; $P(Z_i/\theta_i)$ and you have calculated $P(Z_i)$

i.e.,

$$P(\theta_i/Z_i) = \frac{P(Z_i/\theta_i) P(\theta_i)}{P(Z_i)}$$

With the above general description of the various aspects of statistical decision theory in mind we may look at the decision making process as a game. The two protagonists are the decision maker and nature (or the unknown elements). The game is played thusly: first, the decision maker defines the possible courses of action available to him, he gives his a-priori probabilities as to what is the exact state of nature and he performs an experiment of his choice (which may be not to experiment at all). Then it is nature's turn to go. It chooses an outcome to the experiment, say Z_i according to the probability $P(Z_i)$. Next the decision maker takes the spotlight. He chooses a course of action a_i , which in general will depend on what nature chooses as the outcome Z_i . The next move is up to nature, who chooses a state of nature θ_i according to the probability $P(\theta_i/Z_i)$, the a-posteriori probability conditioned.

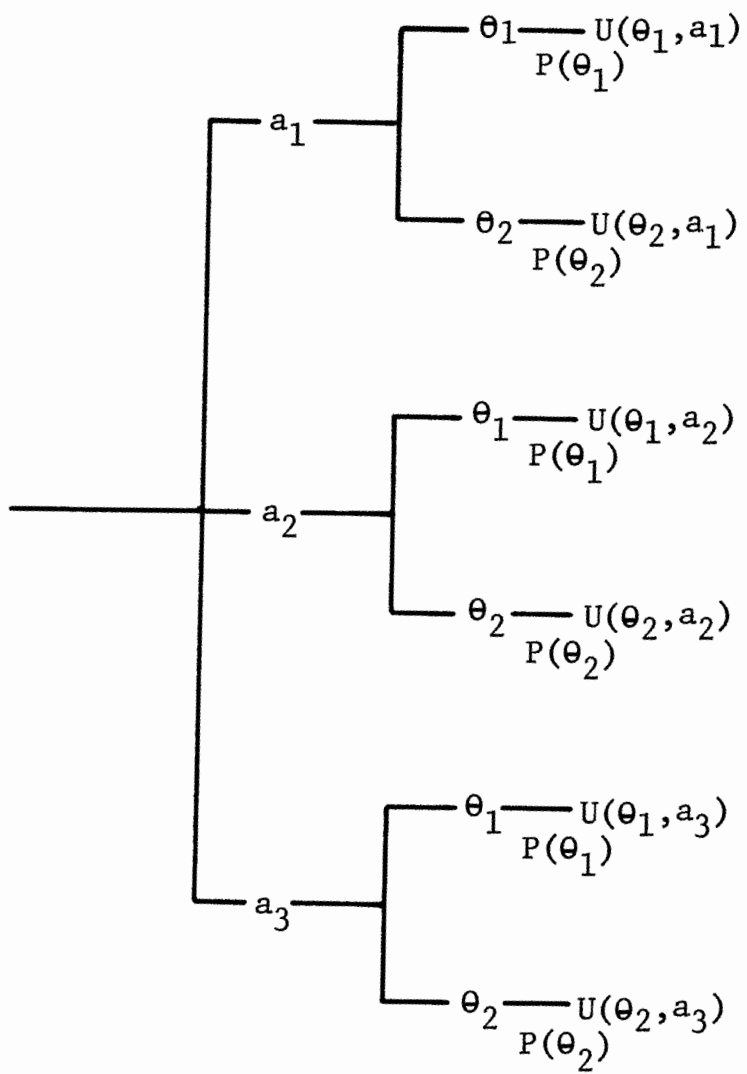
The game is now completed, and now the decision maker will receive his "payoff" or the utility he will receive for performing the experiment and getting the outcome Z_i , taking the action a_i with the exact state of nature being θ_i .

The game or decision making procedure may be characterized by a decision tree such as that shown in Fig. 1. Here we have the nodes of the tree as the points in the process where decisions have to be made. The convention that I will be using is to represent decision nodes at which the decision maker must decide by an X and to represent decision nodes at which nature must decide by a dot (.).

The branches flowing out of a nature decision node are chosen by nature statistically and these statistics are entered below the corresponding branch. The terminal utilities from a utility table are entered after the terminal nodes of our decision tree and the expected utilities are entered above the node corresponding to nature's decision and the expected utility due to the decision maker's decision (which will be oriented to maximize the expected utility) is placed above that corresponding node.

Let us now proceed to look at some of the problems involved in applying the above theory to making legal decisions.

FIG. 1



CHAPTER II
THE PROBLEMS

In the previous chapter we discussed the theory of making decisions utilizing a statistical approach. Well, the problem that is under investigation here is whether or not this statistical decision theory is applicable to making the decisions that arise in the legal profession. As is the case with most investigations of this type the major problem arises in trying to fit a precise and mathematically-defined model, like the one presented in Chapter I, to the harsh realities of the real world. You must be able to associate the proper areas of the real world with the various defined segments of your model.

The procedure that I employed in this thesis is that I interviewed a number of well-reputed attorneys and explained to them what statistical decision theory was, and that I was trying to find some application of this theory to making legal decisions. I then asked them if they would look into their files and come up with some cases in which they thought this theory might apply. We then proceeded to analyze these cases where the attorneys attempted to educate me as to the law and facts involved.

We then selected a few cases which I thought would be especially illustrative for this investigation. I then proceeded to inquire as to what were the attorney's decisions in each of these cases. What alternatives did the attorney see

open to him? What were the unknown factors involved that the attorney had to consider before choosing amongst these several alternatives? Here we drew the analogy between the alternatives of the real world situation and the courses of action a_i of the model. The problem here is that any complex legal decision will usually have a large number of alternatives open to the attorney which, in the face of a large number of states of nature, could lead to a very extensive decision tree, and thus may need a computer to calculate and keep track of all the probabilistic data. Therefore, to keep the calculation within bounds and to provide a better illustrative example, I had to restrict the number of possible courses of action, and in the law case that was finally selected for illustration of this investigation we have only two possible alternatives from which the attorney is to make his decision.

The unknown factors of the real world, mentioned above, are analogous to the states of nature θ_i in our model. The problem here is again that for a very large number of states of nature, the decision tree becomes very complex and more complex computational means must be employed. In addition, I have discovered that a large number of legal decisions must be made in the face of a continuous distribution of states of nature as compared with discrete states of nature as illustrated in this thesis. A continuous distribution of states of nature needs continuous probability theory to describe its statistics and my suggestions on this topic are presented in Chapter V of this text.

Once we have determined what in the real world we are to call the courses of action a_i and what we are to call the states of nature, we need to know the attorney's a-priori probability of the states of nature; i.e., his degree of uncertainty. The problem here is that most attorneys are uncertain about what they are uncertain about. In short, it may be almost impossible to get an attorney to give you the a-priori probabilities.

Also, at this point, in addition to the statistics on the states of nature, we need to ascertain from the attorney just what he expects to gain or lose in terms of utility expressed in relative units when he takes a course of action and finds a state of nature. This information is his utility table discussed in the preceding chapter, and it has as its counterpart in the real world how the attorney feels he has gained or lost either monetarily or in any other way when he decides on a particular alternative and later finds out the true factors involved. In interviewing these attorneys, I have found it very difficult to obtain these utility tables, and I believe the reason is that an attorney is simply not used to the idea of thinking in terms of utilities. Or, at least, he does so only subconsciously. When the states of nature are continuous, then the utilities are no longer discrete entries into a table, but are now a function of the continuous variable θ and have a_i as a parameter. Thus the expected utility is no longer a weighted sum but it is an integral.

The next problem that arises in applying statistical decision theory to legal decision is choosing the proper experiment to investigate. The experiment to increase one's knowledge at a price of the state of nature usually does not come to the attorney's door. The mountain must go to Mohammed, and therein is the problem. If good experiments, and I mean good in the sense of raising the expected utilities by more than it costs, are to be found and analyzed as good, they must be suggested by someone, and since an attorney's time and knowledge are his product, it may turn out to be too expensive to analyze all possible experiments in search of the good ones. A further discussion of the need for a criteria to limit the number of experiments that are worth investigating is presented in Chapter V.

Once an experiment is under analysis, we must obtain from the attorney his estimates of the $P(Z_i/\theta_j)$ or in words, the probability that the experiment will have an outcome Z_i if it is known beforehand that the state of nature is θ_j . This is usually gotten from historical background of the experiment. The problem arises when this historical data is unavailable.

With the above problems in mind let us now analyze a theoretical case in which we apply the statistical decision theory to making a hypothetical legal decision.

CHAPTER III
A THEORETICAL EXAMPLE

With the foregoing concepts in mind, let us now consider a theoretical situation that a typical lawyer might have to encounter. The facts of this case are purely fictitious but the methods used here would be the same as the methods used to apply statistical decision theory in an actual law case. In subsequent chapters we will analyze actual law cases and see how well the theoretical application of this chapter holds for the practical world of the practicing attorney.

In this theoretical example we will analyze the facts of the case and determine what the possible states of nature are. Then we will consider what are the possible courses of action open to our friend the attorney, and for each action we will determine the utility gained or lost by the attorney in light of a particular prevailing state of nature. We then determine the a-priori probabilities and the a-posteriori probabilities of which is the prevailing state of nature. We can then plan our strategies; i.e., our course of action given a particular outcome of our experiment and calculate the expected utility for each of these strategies. Then, based on these statistics, we will decide on the best strategy leading to the maximum utility.

The facts of this case are as follows. A young man, age 23, came into the State Street office of the

world-renowned criminal attorney, the famous Mr. Justice. The young man sat down in front of Mr. Justice's huge mahogany desk and proceeded to speak, "I need your help Mr. Justice. My name is Arnold Friend and I am here to ask you to defend my best buddy who is now in county jail awaiting trial. His name is John Trouble and he is being charged with the murder of Bill Bullet who was found shot to death in his Back-bay apartment."

The first decision that our famous lawyer has to face is whether to handle this case or not. Let us call these two available courses of action a_1 to defend Mr. Trouble, and a_2 not to defend Mr. Trouble. Now, before Mr. Justice has had a chance to do any investigation into the facts of the case, he knows that either Mr. Trouble is guilty or he is innocent, and since Mr. Justice has no further information than the name of the visitor and the name of the accused, he arbitrarily assigns equal probabilities to these states of nature. In other words, the two states of nature are θ_1 , Mr. Trouble is innocent of the crime, or θ_2 , Mr. Trouble is guilty of the crime.

Mr. Justice then thinks to himself, "If I defend Mr. Trouble (a_1) and he is really innocent (θ_1), then I have a good chance to further my career, and I'll earn some money to boot." This seems like the best situation for Mr. Justice and he arbitrarily assigns it a gain of 3 units. This is an arbitrary assignment and it is customary to assign a maximum gain of utility to the best combination of θ_i and a_i and then any other combination (which would be less favorable to Mr.

Justice) would have a gain, less than the maximum, assigned to that θ_i and a_i . "So", thinks Mr. Justice, "if I do not defend Mr. Trouble (a_2) and he is really innocent (θ_1) then I am not doing my part for justice and also I will not receive a salary. However, my reputation cannot be harmed very much; I think I would gain 1 unit relative to the situation of ($a_1\theta_1$) above." Considering the other possibilities, he decides that if he does not defend Mr. Trouble (a_2) and Mr. Trouble is really guilty (θ_2) then his prestige will gain and he gives this gain an assignment of 2 units. Finally, if he does defend Mr. Trouble (a_1) and Mr. Trouble is guilty (θ_2), then his reputation will suffer greatly and his gain is 0 units on the relative scale. Mr. Justice's utilities are summarized in Table I.

After mentally calculating his utility table 1, Mr. Justice whipped out a piece of paper and a pencil and drew the decision tree illustrated in Fig. 2. In this tree our decision making lawyer denotes nodes at which he is faced with a decision by an X and he denotes those nodes where nature makes her decision by a dot (.). The branches of the tree extending from a nature decision node have a symbol above the branch representing which choice that branch corresponds to, and below each branch is the probability with which nature chooses that branch. At the tip of each branch is placed the utility corresponding to traveling that far along the decision tree and next to the nature node is placed the expected utility

TABLE I

	a_1	a_2
θ_1	3	1
θ_2	0	2

TABLE Ia

	a_1	a_2
θ_1	3	1
θ_2	-1	2

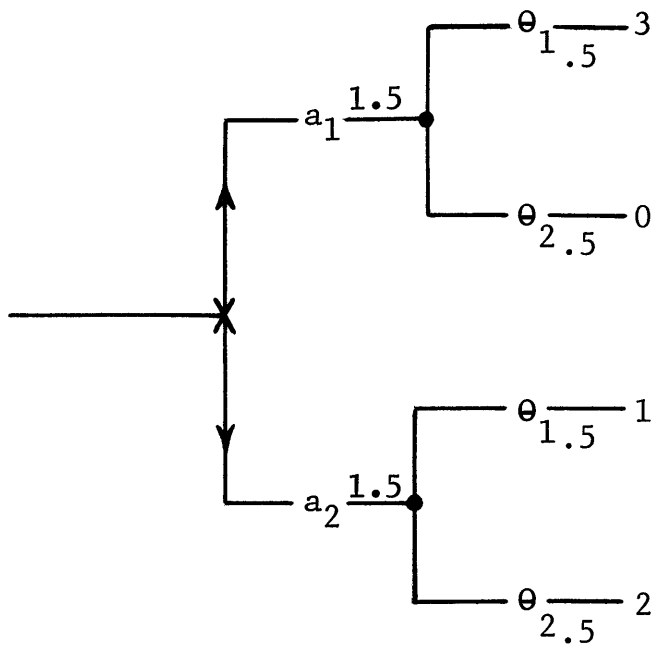


FIG. 2

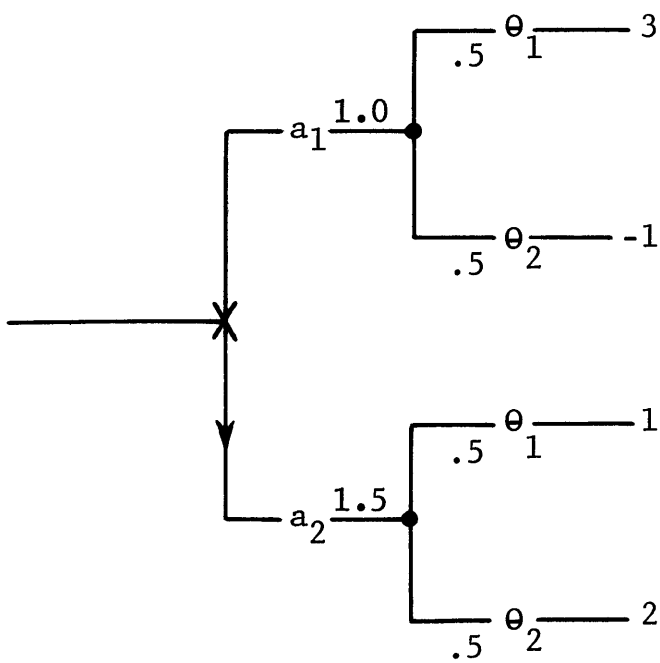


FIG. 2a

for taking the course of action which leads to that nature decision node. Our decision maker, Mr. Justice, will decide on that course of action which will maximize his expected gain.

In this case, with his limited information, Mr. Justice calculates that if he chooses to defend Mr. Trouble, that Mr. Trouble will be innocent with a probability .5 and this will yield a gain of 3 and that Mr. Trouble will be guilty with a probability .5 and this will yield a gain of 0. Therefore, the expected gain for Mr. Justice by taking action a_1 is $(.5)(3) + (.5)(0) = 1.5$. On the other hand, if he decides not to defend Mr. Trouble (a_2), then Mr. Trouble will be innocent (θ_1) with a probability .5 and this will yield a gain of 1 and that Mr. Trouble will be guilty (θ_2) with a probability .5 and this will yield a gain of 2. Therefore, Mr. Justice calculates that by taking action a_2 his expected loss is $(.5)(1) + (.5)(2) = 1.5$.

In this case the expected gain for each course of action a_1 or a_2 is the same and it does not matter which course of action Mr. Justice takes to maximize his expected gain. Even though this is a theoretical example, it is still a model of a true life situation, and as such, the utility table should reflect the true bias of the decision maker. In this case of equal probable $P(\theta_1)$, the decision maker is unbiased as to which course of action to choose. We can then say that Mr. Justice is unbiased as to the exact state of nature. However,

consider what a table such as Table 1a might reflect as to Mr. Justice's bias. In this table 1a only the entry for (a_1, θ_2) is changed and it is decreased from 0 to -1; i.e., Mr. Justice is now more sensitive to a loss of reputation (maybe this is an election year and Mr. Justice is running for D.A.). The decision tree now takes on the shape of Fig. 2a, and mirroring the above calculation, Mr. Justice calculates that his expected gain of utility for taking action a_1 is 1, and for taking action a_2 is still 1.5. Now, Mr. Justice is advised to select action a_2 to maximize his expected gain. We can now say that Table 1a reflects a_2 . And similar analyses changing other entries in Table I might lead to choosing action a_1 with equal $P(\theta_1)$ and thus reflect a decision maker biased toward action a_1 .

Using Table 1 as Mr. Justice's utilities, we found that it made no difference to him which action he chose with regard to maximizing his expected gain.

However, the assumed $P(\theta_1)$ being equal-probable represents the fact that, as astute as Mr. Justice may be, he has absolutely no knowledge as to whether Mr. Trouble is guilty or innocent. Of course, since Mr. Justice is so astute, he now considers the possibilities of obtaining more information about the guilt or innocence of Mr. Trouble by paying a price. In other words, through experimentation (at a cost), Mr. Justice hopes to improve his knowledge of the true state of nature and then by choosing an appropriate

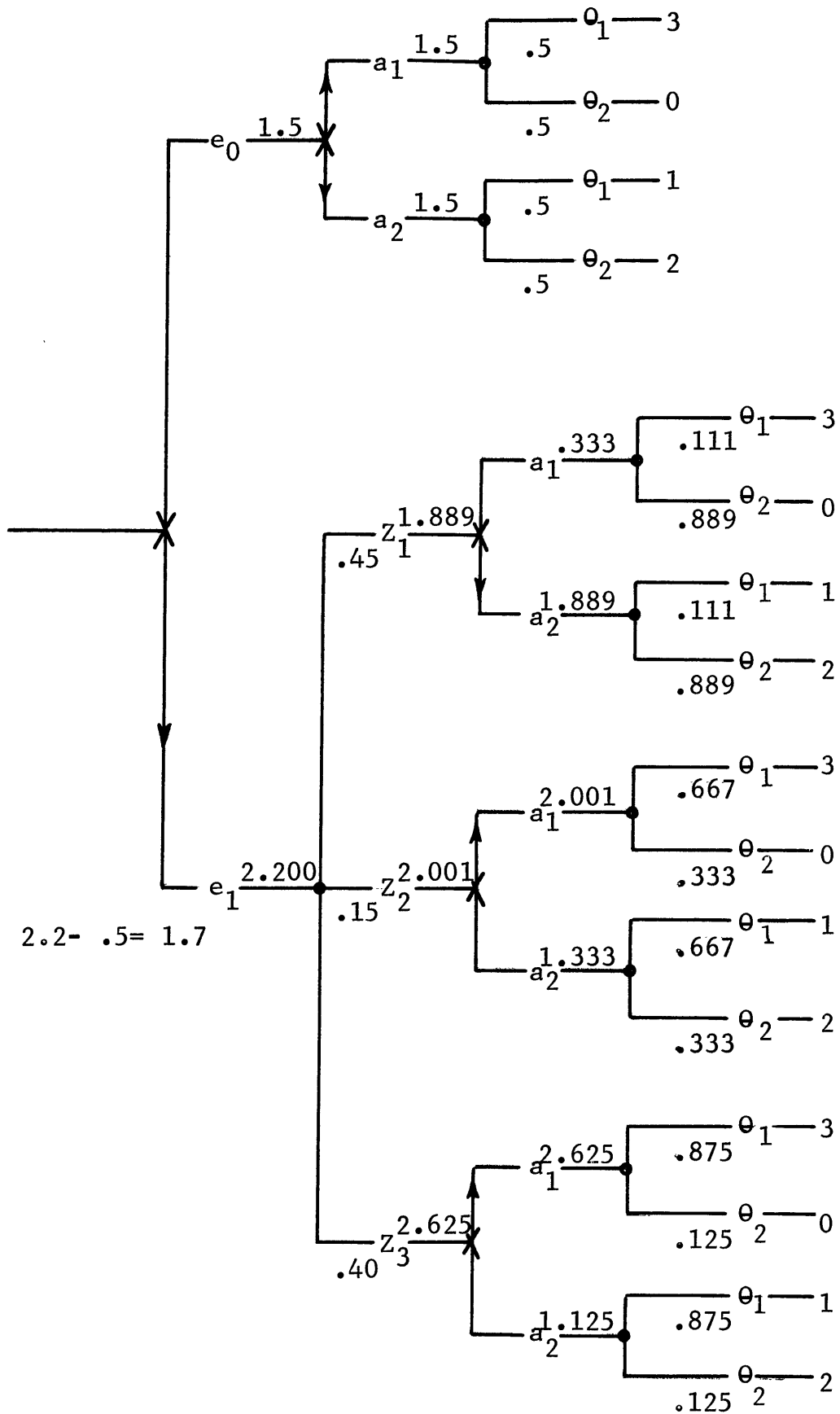
course of action, maximize the expected gain in utility still further. Before we consider the introduction of an experiment, let us ponder the value to Mr. Justice of information which would tell him exactly what the prevailing state of nature is. For example, if he were told that Mr. Trouble is innocent, he would defend him for a gain of 3. If he were told that Mr. Trouble were guilty, then he would not defend him for a gain of 2. At this point in the game, Mr. Justice would be told this perfect information with the a-priori probabilities. In other words, he would be told that Mr. Trouble was innocent with a probability of .5 and he would be told that Mr. Trouble was guilty with a probability of .5. Therefore, if Mr. Justice were given this perfect information he would expect a gain of $.5(3) + .5(2) = 2.5$. Before he had this perfect information he could take either action a_1 or a_2 and expect a gain of 1.5. The perfect information represents a gain to Mr. Justice of 1 unit and he should be willing to pay any price up to 1 unit for it. Also, since any experiment he can hope to perform to obtain a better knowledge of the state of nature can give him only information inferior to the perfect information, 1 unit represents a maximum he should be willing to expend for any experiment.

With this in mind, Mr. Justice considers making use of his life-long friend, the eminent private investigator Mr. Holmes. He has used Mr. Holmes in the past and he knows that Mr. Holmes will charge him an amount which, measured in utility, will be .5 units. For this sum Mr. Holmes

will make an investigation of the case and will report back to Mr. Justice one of three reports. These three possible reports we will call Z_1 , Z_2 and Z_3 . Z_1 represents the report that Mr. Trouble is "very guilty" Z_2 represents the report that Mr. Holmes does not know whether Mr. Trouble is guilty or not even after the investigation. Finally, Z_3 represents the report that Mr. Trouble is "very innocent".

Using the same notation as above we draw the decision tree illustrated in Fig. 3. The first decision that Mr. Justice must now make is whether to have Mr. Holmes investigate (represented by e_1) or to make a decision without any further information (represented by e_0). If Mr. Justice decides not to have an investigation, then the decision tree branching off of e_0 will be the same as Fig. 2. If, however, Mr. Justice decides to make use of Mr. Holmes' services, he will decide e_1 . The next decision will be nature's and she will decide on one of the three reports Z_1 , Z_2 , Z_3 with corresponding probabilities $P(Z_1)$, $P(Z_2)$, $P(Z_3)$. These probabilities are calculated by Mr. Justice in the following way. First, he realizes that based upon his many years of using the services of Mr. Holmes that if Mr. Trouble is innocent, Mr. Holmes will report Z_1 with probability .1 (i.e., $P(Z_1/\theta_1) = .1$), Z_2 with probability .2 (i.e., $P(Z_2/\theta_1) = .2$), and Z_3 with probability .7 (i.e., $P(Z_3/\theta_1) = .7$). He also realizes that if Mr. Trouble is guilty that Mr. Holmes will report Z_1 with probability .8 (i.e., $P(Z_1/\theta_2) = .8$), Z_2 with

FIG. 3



probability .1 (i.e., $P(Z_2/\theta_2) = .1$), and Z_3 with probability .1 (i.e., $P(Z_3/\theta_2) = .1$). From these probabilities and using the equation from Appendix, we have:

$$\begin{aligned}
 P(Z_1) &= P(Z_1/\theta_1) P(\theta_1) + P(Z_1/\theta_2) P(\theta_2) \\
 &= (.1) (.5) + (.8) (.5) \\
 P(Z_1) &= .45
 \end{aligned}$$

$$\begin{aligned}
 P(Z_2) &= P(Z_2/\theta_1) P(\theta_1) + P(Z_2/\theta_2) P(\theta_2) \\
 &= (.2) (.5) + (.1) (.5) \\
 P(Z_2) &= .15
 \end{aligned}$$

$$\begin{aligned}
 P(Z_3) &= P(Z_3/\theta_1) P(\theta_1) + P(Z_3/\theta_2) P(\theta_2) \\
 &= (.7) (.5) + (.1) (.5) \\
 P(Z_3) &= .40
 \end{aligned}$$

These probabilities are entered under the appropriate branches.

Now that Mr. Justice has $P(Z_i/\theta_i)$, $P(\theta_i)$ and $P(Z_i)$ in his arsenal of statistics, he now calculates the a-posteriori probabilities from the Bayesian relationship:

$$P(\theta_i/Z_i) = \frac{P(Z_i/\theta_i) P(\theta_i)}{P(Z_i)}$$

$$(1) \quad P(\theta_1/Z_1) = \frac{P(Z_1/\theta_1) P(\theta_1)}{P(Z_1)} = \frac{(.1)(.5)}{.45} = .111$$

$$(2) \quad P(\theta_2/Z_1) = \frac{P(Z_1/\theta_2) P(\theta_2)}{P(Z_1)} = \frac{(.8)(.5)}{.45} = .889$$

$$(3) \quad P(\theta_1/Z_2) = \frac{P(Z_2/\theta_1) P(\theta_1)}{P(Z_2)} = \frac{(.2)(.5)}{.15} = .667$$

$$(4) \quad P(\theta_2/Z_2) = \frac{P(Z_2/\theta_2) P(\theta_2)}{P(Z_2)} = \frac{(.1)(.5)}{.15} = .333$$

$$(5) \quad P(\theta_1/Z_3) = \frac{P(Z_3/\theta_1) P(\theta_1)}{P(Z_3)} = \frac{(.7)(.5)}{.40} = .875$$

$$(6) \quad P(\theta_2/Z_3) = \frac{P(Z_3/\theta_2) P(\theta_2)}{P(Z_3)} = \frac{(.1)(.5)}{.40} = .125$$

After nature chooses the report, Mr. Justice then chooses between a_1 and a_2 such that his expected gain in utility is maximum and this expected gain is calculated as above. For example, if nature were to choose Z_1 , then Mr. Justice could choose a_1 and nature could choose θ_1 , with probability $P(\theta_1/Z_1) = .111$ and yield a utility of 3, or nature could choose θ_2 with probability $P(\theta_2/Z_1) = .889$ and yield a utility of 0. Thus the expected gain in utility to Mr. Justice for choosing a_1 when nature has chosen Z_1 is $(.111)(3) + (.889)(0) = .333$. Similarly, if Mr. Justice had chosen a_2 then nature could choose θ_1 with probability $P(\theta_1/Z_1) = .111$ and yield a utility of 1, or nature could choose θ_2 with

probability $P(\theta_2/Z_1) = .889$ and yield a utility of 2. Thus the expected gain in utility to Mr. Justice for choosing a_2 when nature has chosen Z_1 is $(.111)(1) + (.889)(2) = 1.889$. Thus wishing to maximize his utility gain, Mr. Justice would choose a_2 if nature chose Z_1 and he would expect to gain 1.889. If nature were to choose Z_2 , then Mr. Justice could choose a_1 and nature, in turn, could choose θ_1 , with probability $P(\theta_1/Z_2) = .667$ and yield a utility of 3, or nature could choose θ_2 with probability $P(\theta_2/Z_2) = .333$ and yield a utility of 0. Thus the expected gain in utility to Mr. Justice for choosing a_1 when nature has chosen Z_2 is:

$$(.667)(3) + (.333)(0) = 2.001$$

Similarly, if Mr. Justice had chosen a_2 , then nature could choose θ_1 with probability $P(\theta_1/Z_2) = .667$ and yield a utility of 1, or nature could choose θ_2 with probability $P(\theta_2/Z_2) = .333$ and yield a utility of 2. Thus the expected gain in utility to Mr. Justice for choosing a_2 when nature has chosen Z_2 is:

$$(.667)(1) + (.333)(2) = 1.333$$

Thus wishing to maximize his utility gain, Mr. Justice would choose a_1 if nature chose Z_2 and he would expect to gain 2.001.

By analogous calculations, if nature chooses Z_3 , then Mr. Justice will expect a gain in utility of 2.625 if he chooses a_1 , and a gain of 1.125 if he chooses a_2 . Therefore, to

maximize his expected gain he will choose a_1 and expect to gain 2.625.

So we see that when Mr. Justice chooses to have Mr. Holmes investigate (e_1) nature chooses Z_1 with probability $P(Z_1) = .45$, and Mr. Justice can expect to gain 1.889, or nature chooses Z_2 with probability $P(Z_2) = .15$ and Mr. Justice can expect to gain 2.001 or finally nature chooses Z_3 with probability $P(Z_3) = .40$ and Mr. Justice can expect to gain 2.625. Thus, the expected gain in utility to Mr. Justice for having Mr. Holmes investigate is $(.45)(1.889) + (.15)(2.001) + (.40)(2.625) = 2.200$. However, Mr. Justice must pay Mr. Holmes .5 for his results and therefore the net expected gain in utility due to Mr. Holmes' investigation is $2.200 - .500 = 1.700$ and this is entered under branch e_1 .

This gain in utility of 1.7 due to Mr. Holmes' investigation is greater than the expected gain in utility when no investigation is made. Therefore, Mr. Justice would be very well advised to have Mr. Holmes perform his investigation. In fact, if Mr. Holmes had been a statistical decision maker, he could have charged up to .7 units of utility and Mr. Justice would still buy.

CHAPTER IV.

A Practical Example

This chapter is devoted to testing if the model and methods utilized in Chapter III to make synthetic legal decisions can be applied to making non-synthetic legal decisions in the analysis of an actual law case. The following analysis represents several months of interviews with some of the top legal talents in the Boston area. Together with these distinguished attorneys, I have discussed many of the facets to making legal decisions and have tried to isolate these decisions for application of the above-mentioned decision theory model.

The approach that I found best in this process of interviewing the legal community was to explain just what statistical decision theory was, without being over-technical, to give some very simple applications of the theory to making very elementary decisions. After this introductory process was concluded, the attorney delved into his file for fairly current cases that might be particularly apropos for this investigation. The reason that I specifically wanted a "fairly current" case was that most attorneys handle a large number of cases over a period of time, say a year, and to be able to answer my query "What were the decisions you had to make in this case?", the case had to be still alive in his mind.

Ideally, I selected cases for this thesis in which decisions were still being made at this writing.

Once an appropriate case had been chosen or, more precisely, once a particular decision in an appropriate case had been chosen for investigation, the next question was, "What were the various alternatives of action available to you?", and "Under what possible uncertainties in the state of the world were you laboring under which prevented you from choosing a particular alternative of action with absolute certainty?" In terms of Chapter III, I was asking what were the possible courses of action a_i and what were the possible states of nature θ_i . Isolating the possible courses of action was not as difficult as defining what were the possible states of nature and in many cases that were brought out of the attorney's file, it was impossible to define the possible states of nature.

However, once a case was found, in which a decision and the various prevailing states of nature under which this decision had to be made were well defined, I then had to determine from the attorney what were his a-priori feelings as to the relative likelihoods of each of the possible states of nature. This was usually a process comparable to pulling teeth, because most attorneys do not have a specific feel for what you mean by the "probability of a particular state of nature" and, therefore, you must be very careful to avoid the word "probability" and talk about the more familiar things

such as "from your experience and under the present circumstance, what is the percent of the time you would expect the state of nature to be such and such over a long period of observation?" Phrasing the query in this way has the effect of not alienating or confusing the attorney by using terms which may be conceptually unknown to him and also putting the question in terms that may be closer to what he was actually thinking when he was faced with the decision.

Once I successfully reached this far in my analysis, I tended a sigh of relief, but progress was short-lived however, since my next query usually dampened the party again. This was the attorney's utility table. I found that if I came right out and asked the attorney what was his utility for taking a particular course of action, and finding that a particular state of nature was prevailing, that I usually got a blank stare as a response. This was because the attorney invariably interpreted this to mean, what would be the absolute gain to him if he took a course of action and the state of nature was such and such. This would be almost impossible to judge especially if the utility was not measured in money alone but included such things as speed of trial, prestige, etc.

However, an absolute number representing these utilities is not what is needed to make a decision. What is really needed is a scale usually in the lawyer's own mind, based on years of experience, which measures the relative value

of taking a particular course of action say a_1 and having a particular state of nature θ_1 , with taking a particular course of action a_i and having a particular state of nature θ_i . This is many times easier for an attorney to estimate than the absolute scale referred to above. So, my query to determine the utility of the attorney concerning the decision under investigation was usually this: "Consider all the possible courses of action available to you and all the possible states of nature. List all the possible combinations of a course of action with a state of nature. Then rearrange this list so that the combination which is the best for you is at the top of the list, and that combination which is the worst for you is at the bottom of the list. Then assign a number to each combination such that the number you assign to one combination (of course of action and state of nature) will reflect how much better for you this combination is, or how much worse for you it is than some other combination".

Once these relative utilities have been ascertained along with the statistical experience of the attorney about the possible states of nature, we were ready to approach the decision between the various courses of action. This entailed drawing an elementary tree such as the one discussed in Chapter I, and depicted in Fig. 1. Also at this point, I was able to discuss with the attorney the possible ways in which he might proceed to get more information about the existing state of nature so that he could increase his

expected gain in utility by the decision making process. We then analyzed these possible experiments according to the theory outlined in Chapter I and put to test in the example of Chapter III.

This essentially is the format that the interviews usually took. Let us now look at the facts of the specific case now before us. This case concerns an automobile accident which occurred in the summer of 1961. The attorney whom I interviewed is handling the case for the plaintiff, a boy who, in the summer of 1961, was getting ready to enter his senior year in high school. The year before, this lad was awarded the Harvard Book award for scholastic achievement, and although a quiet boy, he partook in many high school activities and was quite friendly. In short, this boy has an excellent scholastic and extra-curricular record thus far in his high school career.

Then in the summer of 1961 this boy was a passenger in an automobile driven by his friend. There were three boys in this car, and according to witnesses, the boys were not speeding. Then a woman driving down a cross street, failed to stop her motor vehicle in time and plowed straight into the right front door of the boys' car behind which the plaintiff was sitting as passenger. The plaintiff, as well as other occupants, was thrown from the car, and the plaintiff was in a coma for nineteen days and suffered a fractured pelvis, a fractured jaw and several fractured ribs. In

addition, he had multiple skull fractures and lacerations of the brain with brain fluid coming out of one ear and both nostrils. When he arrived at the hospital the admitting surgeon did not give him much of a chance to live.

He did live, however, and after much medical treatment he became medically well. However, he was not mentally well. He suffered a complete change in his personality. When he finally went back to his senior year in high school, he could not get along well with anyone. He was completely withdrawn. He also did very poorly scholastically; in fact, he flunked several subjects and could not get into any college to which he applied. He finally got into a second-rate college only after a year of prep school and is doing only "C" work there. In short, what promised to be such a good career has been completely thwarted.

At this point, I asked the attorney what type of decision did he have to make in this case. He replied that the most pressing and most basic decision in this case is whether to build up as large a case as possible for the plaintiff by having extensive psychiatric testimony and medical testimony as well as the testimony of friends and teachers who knew this boy both before and after the accident, and all of whom could testify to the boy's change in personality. This would require a substantial expenditure of money. Or, would it be better to rely on the medical record so far and the case as it now stands and have not

as big a case, but not expending a lot of money to build it?

Let us then call these two courses of action a_1 , build up a big case and a_2 , build up a small case. The former course of action will cost an estimated \$3,000, according to the attorney. At this point, I tried to determine the possible states of nature, and I asked the attorney what were the possible states of the world that would determine which course of action he would take. He replied that it all depended on whether the defendant has a sufficient amount of insurance coverage to pay a large judgment if they were to press for one. If she had only the minimum coverage (\$5,000/10,000) and they pressed for a large case, all they would collect is a few thousand dollars and this would be expended in building up a big case.

So, for simplicity, let us assume that there are three possible states of nature, θ_1 she has only minimum coverage, θ_2 she has moderate coverage and θ_3 she has adequate coverage for any judgment against her. Obviously, I could have formed many more states of nature, in fact, one for each possible insurance level of coverage she could have been carrying. However, for the present illustrative purposes we will use only three states of nature, and be aware of the possible extension to more.

I asked the attorney if this information about the defendant's insurance coverage might not be easily obtainable. He answered that it could not. That one of his

partners who is an officer in the council of Boston Insurance Companies could not even obtain this information. At this point, I tried to ascertain what were the attorney's feelings as to the relative likelihood of these three states of nature. He felt that because the defendant was of limited means, single and owns a modest single dwelling apartment (less than \$10,000) that probably θ_1 was the state of nature. However, since the defendant's insurance company has not notified the attorney that the defendant has only minimum coverage, the attorney feels he must give some weight to θ_2 . He feels also that θ_3 is unlikely. In the light of this, I suggested the following a-priori probabilities and he agreed.

$$P(\theta_1) = .7; \quad P(\theta_2) = .25; \quad P(\theta_3) = .05$$

Then I asked the attorney about the net monetary gain for taking a particular course of action when a particular state of nature existed. He replied that if he "sued big" (a_1) the plaintiff would probably gain \$2,000, if the defendant had minimum coverage (θ_1); would probably gain \$12,000, if the defendant had moderate coverage (θ_2) and would probably gain \$47,000, if the defendant had adequate coverage (θ_3). On the other hand, if he "sued small" (a_2) the plaintiff would probably gain \$5,000, if the defendant had minimum coverage (θ_1); would probably gain \$15,000, if

the defendant had moderate coverage θ_2 , and would probably gain \$15,000., if the defendant had adequate coverage θ_3 . These utilities are summarized in Table 2.

Let us now apply the model discussed in Chapter III and construct a decision tree similar to the tree in Fig. 2. This tree is shown in Fig. 4. The first decision is up to the lawyer and he will choose between the two courses of action "sue big" (a_1) and "sue small" (a_2). Then nature takes her turn in this decision game and she will choose among the three possible states of nature with a-priori probability:

$$P(\theta_1) = .7; \quad P(\theta_2) = .25 \text{ and } P(\theta_3) = .05$$

These are written below the appropriate branch coming from nature's decision nodes. The utilities are placed at the end of these branches and represent the net gain of the plaintiff for choosing the course of action and finding that state of nature which leads him along that branch.

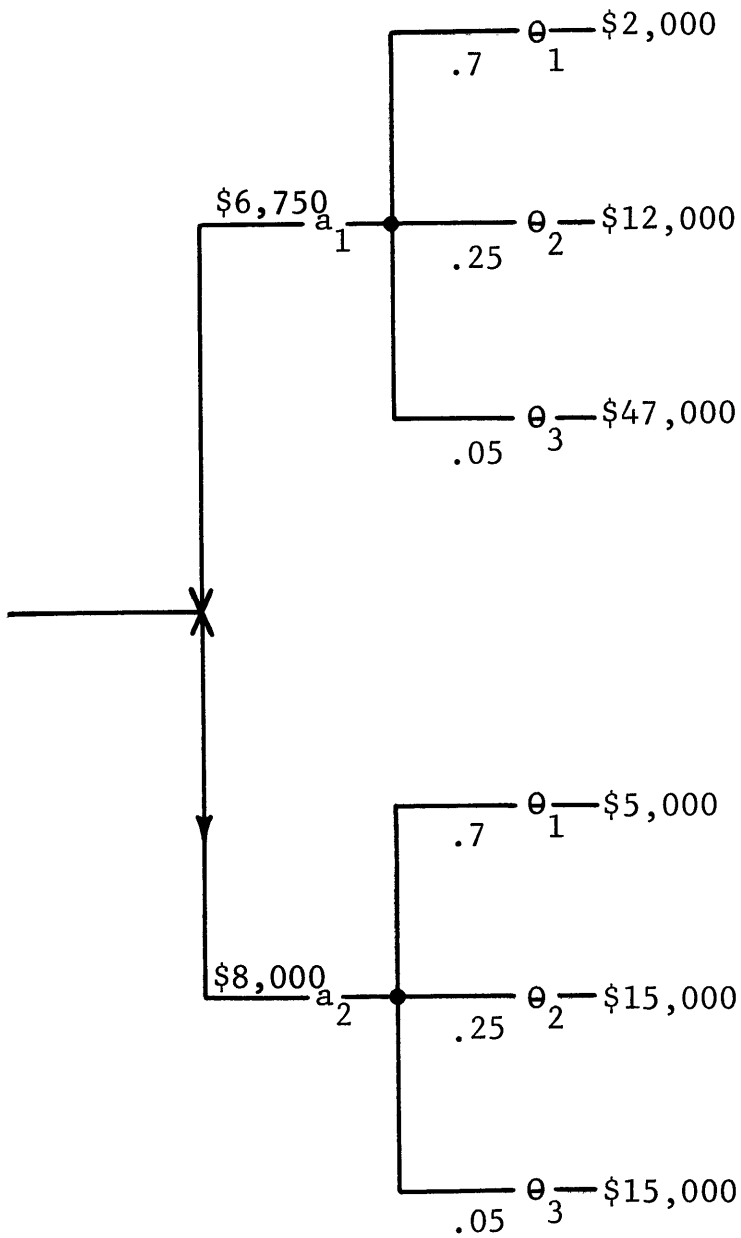
We can calculate the expected gain in utility for each course of action as we did in Chapter III. If the attorney chooses to "sue big" (a_1), then nature may choose θ_1 with probability = .7 and the gain = \$2,000., or nature may choose θ_2 with probability = .25 and the gain = \$12,000., or nature may choose θ_3 with probability = .05 and the gain = \$47,000. Thus the expected gain in utility for "suing big"

TABLE 2

	a_1	a_2	$P(\theta_i)$
θ_1	\$2,000	\$5,000	.7
θ_2	\$12,000	\$15,000	.25
θ_3	\$47,000	\$15,000	.05

$U(\theta_i, a_j)$

FIG. 4



is $.7 (\$2,000.) + .25 (\$12,000.) + .05 (\$47,000.) = \$1,400. + \$3,000. + \$2,350. = \$6,750.$ However, if the attorney chooses to "sue small" (a_2), then nature may choose θ_1 with probability = .7, and the gain = \$5,000; or nature may choose θ_2 with probability = .25 and the gain = \$15,000.; or nature may choose θ_3 with probability = .05 and the gain = \$15,000. Thus the expected gain in utility for "suing small" is:

$$.7(\$5,000.) + .25 (\$15,000.) + .05 (\$15,000) \\ = \$3,500. + \$3,750. + \$750. = \$8,000.$$

Therefore, with these a-priori probabilities, the attorney would be well advised to "sue small" (i.e., take course of action a_2) in order that he maximize his expected gain in utility. In this case, he now can expect to gain \$8,000 by suing small.

Let us now calculate the value of information, which would tell him the exact state of nature. We call the difference between this expected gain in utility before exact knowledge and after the expected value of perfect information, E.V.P.I. It is the value of knowing of the exact state of nature and, as such, represents the maximum amount one should pay for any information which would let you know something about the exact state of nature. If the attorney was told what the exact state of nature is, he would be

told that it is θ_1 with probability $P(\theta_1) = .7$, and he would take action a_2 and gain \$5,000; or, he would be told that it is θ_2 with probability $P(\theta_2) = .25$ and he would take action a_2 and gain \$15,000, or he would be told that it is θ_3 with probability $P(\theta_3) = .05$ and he would choose action a_1 and gain \$47,000. Thus his expected gain in utility for perfect information is:

$$.7 (\$5,000) + .25 (\$15,000) + .05 (\$47,000) \\ = \$3,500. + \$3,750. + \$2,350. = \$9,600.$$

However, before he had exact knowledge he could expect to gain \$8,000. by taking action a_2 . Therefore, the E.V.P.I. is $\$9,600 - \$8,000 = \$1,600$.

One of the beauties of this model is that it can be used as in Chapter III to analyze the worth of any experiment; that is, suggest to give more information as to the exact state of nature. Let us now analyze the following experiment suggested by me to the attorney. I suggested that he hire an insurance investigator to investigate the defendant's financial status and then investigate insurance company statistics and to report back one of three results, Z_1 , the defendant probably carries low insurance; Z_2 , the defendant probably carries moderate insurance, or Z_3 , the defendant probably carries high insurance.

What is now needed is some statistical information on such things as how often the investigator will report that the person he is investigating has low insurance when he actually has moderate or high insurance. This can only be gotten from experience, and since this attorney has never tried this experiment before, we must merely guess at these statistics. In other words, we will assume that if the defendant actually has low insurance that the investigator will report that she has low insurance 80% of the time, and he will report that she has moderate insurance 15% of the time, and he will report she has high insurance 5% of the time. If, however, the defendant actually had moderate insurance coverage, the investigator will report that she has low insurance 20% of the time, and he will report that she has moderate insurance 70% of the time, and he will report that she has high insurance 10% of the time. If the defendant actually had high insurance coverage, the investigator will report that she has low insurance 10% of the time, that she has moderate insurance 20% of the time, and that she has high insurance 70% of the time. These statistics on the investigator's reliability are summarized in Table 3. Let us now calculate the probabilities that nature will choose the outcomes of this experiment. From the Appendix, we have:

TABLE 3

	z_1	z_2	z_3
θ_1	.7	.25	.05
θ_2	.2	.7	.1
θ_3	.1	.2	.7



$P(z_i/\theta_j)$

$$\begin{aligned}
P(Z_1) &= P(Z_1/\theta_1) P(\theta_1) + P(Z_1/\theta_2) P(\theta_2) + P(Z_1/\theta_3) P(\theta_3) \\
&= (.7) (.7) + (.2) (.25) + (.1) (.05) \\
P(Z_1) &= .545
\end{aligned}$$

$$\begin{aligned}
P(Z_2) &= P(Z_2/\theta_1) P(\theta_1) + P(Z_2/\theta_2) P(\theta_2) + P(Z_2/\theta_3) P(\theta_3) \\
&= (.25) (.7) + (.7) (.25) + (.2) (.05) \\
P(Z_2) &= .360
\end{aligned}$$

$$\begin{aligned}
P(Z_3) &= P(Z_3/\theta_1) P(\theta_1) + P(Z_3/\theta_2) P(\theta_2) + P(Z_3/\theta_3) P(\theta_3) \\
&= (.05) (.7) + (.1) (.25) + (.7) (.05) \\
P(Z_3) &= .095
\end{aligned}$$

Now, as we did in Chapter III, we can calculate the a-posteriori probabilities from Bayes' rule:

$$P(\theta_i/Z_j) = \frac{P(Z_j/\theta_i) P(\theta_i)}{P(Z_j)}$$

$$\begin{aligned}
i &= 1, 2, 3 \\
j &= 1, 2, 3
\end{aligned}$$

So:

$$(1) \quad P(\theta_1/Z_1) = \frac{P(Z_1/\theta_1) P(\theta_1)}{P(Z_1)} = \frac{(.7) (.7)}{.545} = .899$$

$$(2) \quad P(\theta_2/Z_1) = \frac{P(Z_1/\theta_2) P(\theta_2)}{P(Z_1)} = \frac{(.2) (.25)}{.545} = .092$$

$$(3) \quad P(\theta_3/z_1) = \frac{P(z_1/\theta_3) P(\theta_3)}{P(z_1)} = \frac{(.1) (.05)}{.545} = .009$$

$$(4) \quad P(\theta_1/z_2) = \frac{P(z_2/\theta_1) P(\theta_1)}{P(z_2)} = \frac{(.25) (.7)}{.360} = .486$$

$$(5) \quad P(\theta_2/z_2) = \frac{P(z_2/\theta_2) P(\theta_2)}{P(z_2)} = \frac{(.7) (.25)}{.360} = .486$$

$$(6) \quad P(\theta_3/z_2) = \frac{P(z_2/\theta_3) P(\theta_3)}{P(z_2)} = \frac{(.2) (.05)}{.360} = .028$$

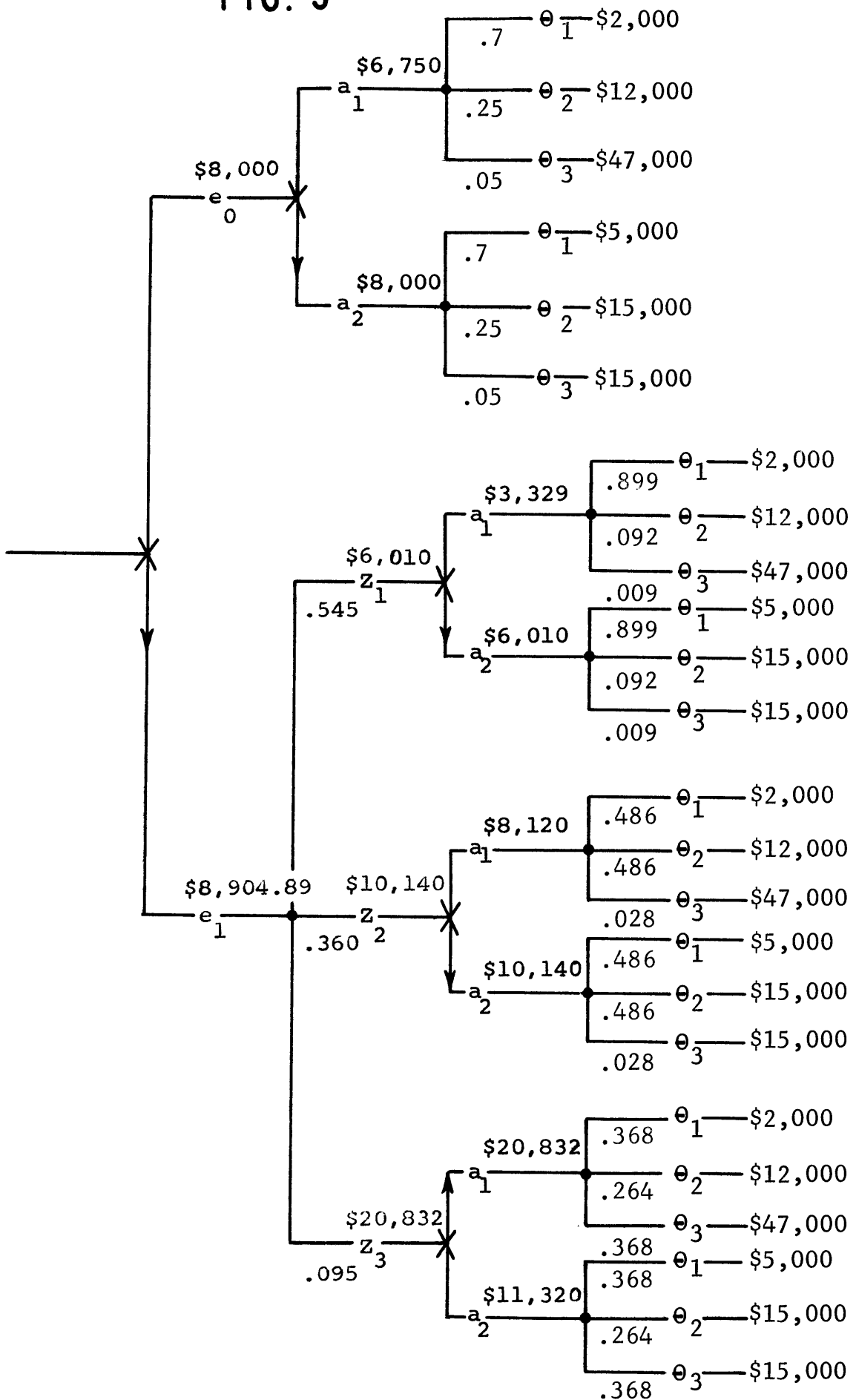
$$(7) \quad P(\theta_1/z_3) = \frac{P(z_3/\theta_1) P(\theta_1)}{P(z_3)} = \frac{(.05) (.7)}{.095} = .368$$

$$(8) \quad P(\theta_2/z_3) = \frac{P(z_3/\theta_2) P(\theta_2)}{P(z_3)} = \frac{(.1) (.25)}{.095} = .264$$

$$(9) \quad P(\theta_3/z_3) = \frac{P(z_3/\theta_3) P(\theta_3)}{P(z_3)} = \frac{(.7) (.05)}{.095} = .368$$

Now we are ready to analyze the decision tree depicted in Fig. 5. The attorney's initial choice is whether to experiment (e_1) or not to experiment (e_0). If he chooses not to experiment he is back to the tree in Fig. 4 and he chooses action a_2 and expects to gain \$8,000. for his client. If, on the other hand, he chooses to experiment, nature then chooses

FIG. 5



one of three possible outcomes with probabilities $P(Z_i)$ $i = 1, 2, 3$. Then it is the attorney's turn to go and he chooses either a_1 or a_2 depending on which course of action will maximize his expected gain in utilities. Finally, nature picks one of three possible states of nature with the a-posteriori probabilities $P(\theta_i/Z_j)$ calculated above (these are placed under the appropriate branches in our tree), and gives the lawyer the utility marked at the end of these branches.

So, let us now proceed to evaluate the experiment by calculating the expected gain in utility if the attorney chooses e_1 . Consider when nature chooses Z_1 as the outcome of the investigation. If the attorney chooses a_1 then nature could choose θ_1 with probability $P(\theta_1/Z_1) = .899$ and yield a gain of \$2,000., or choose θ_2 with probability $P(\theta_2/Z_1) = .092$ and yield a gain of \$12,000, or choose θ_3 with a probability $P(\theta_3/Z_1) = .009$ and yield a gain of \$47,000. Thus the expected gain for the attorney if he chooses a_1 when nature chooses Z_1 is:

$$(.899)(\$2,000) + (.092)(\$12,000) + (.009)(\$47,000) = \$3,329$$

If the attorney chooses a_2 when nature chooses Z_1 then nature could choose θ_1 with probability $P(\theta_1/Z_1) = .899$ and yield \$5,000 or she could choose θ_2 with probability $P(\theta_2/Z_2) = .92$ and yield \$15,000, or she could choose θ_3 with probability $P(\theta_3/Z_1) = .009$ and yield \$15,000. Thus, if the

attorney chooses a_2 when nature chooses Z_1 then his expected gain in utility is:

$$(.899)(\$5,000) + (.092)(\$15,000) + (.009)(\$15,000) = \$6,010.$$

Therefore, wishing to maximize his expected gain in utility, the attorney is advised to choose a_2 if the investigator reports Z_1 , and his expected gain in utility would be \$6,010.

Assume now that nature chooses Z_2 as the outcome of the experiment. If the attorney chooses a_1 then nature could choose θ_1 with probability $P(\theta_1/Z_2) = .486$ and yield \$ 2,000, or choose θ_2 with probability $P(\theta_2/Z_2) = .486$ and yield \$12,000, or choose θ_3 with probability $P(\theta_3/Z_2) = .028$ and yield \$47,000. Thus, the expected gain for the attorney if he chooses a_1 when nature chooses Z_2 is:

$$(.486)(\$2000) + (.486)(\$12000) + (.028)(\$47,000) = \$8,120.$$

If the attorney chooses a_2 then nature could choose θ_1 with probability $P(\theta_1/Z_2) = .486$ and yield \$5,000 or she could choose θ_2 with probability $P(\theta_2/Z_2) = .486$ and yield \$15,000, or choose θ_3 with probability $P(\theta_3/Z_2) = .028$ and yield \$15,000. Thus if the attorney chooses a_2 when nature chooses Z_2 he can expect a gain in utility of:

$$(.466)(\$5000) + (.486)(\$15000) + (.028)(\$15000) = \$10,140.$$

Therefore, wishing to maximize his expected utility, the attorney is well advised to choose a_2 if the outcome of the investigation is Z_2 , and thus expect to gain \$10,140.

Finally, now assume that nature picked Z_3 as the outcome of the investigation. If the attorney chooses a_1 , then nature could choose θ_1 with probability $P(\theta_1/Z_3) = .368$ and yield \$2,000, or she could choose θ_2 with probability $P(\theta_2/Z_3) = .264$ and yield \$12,000, or she could choose θ_3 with probability $P(\theta_3/Z_3) = .368$ and yield \$47,000. Thus, if the attorney chooses a_1 when nature chooses Z_3 he can expect a gain in utility of:

$$(.368)(\$2000) + (.264)(\$12,000) + (.368)(\$47,000) = \$20,832$$

If the attorney chooses a_2 then nature could choose θ_1 with probability $P(\theta_1/Z_3) = .368$ and yield a gain of \$5000. or she could choose θ_2 with probability $P(\theta_2/Z_3) = .264$ and yield \$15,000, or she could pick θ_3 with probability $P(\theta_3/Z_3) = .368$ and yield \$15,000. Thus if the attorney chooses a_2 when nature chooses Z_3 , he can expect a gain in utility of:

$$(.368)(\$5000) + (.264)(\$15,000) + (.368)(\$15,000) = \$11,320$$

In this case, when nature chooses Z_3 as the outcome of the investigation, the attorney still wishing to maximize his expected gain in utility will chose a_1 and expect a gain of \$20,832.

Now we can evaluate the expected gain in utility to the attorney when he chooses to experiment. Nature may choose Z_1 as the outcome with probability $P(Z_1) = .545$, and the attorney chooses a_2 and expects a gain in utility of \$6,010., or nature may choose Z_2 with probability $P(Z_2) = .360$, and the attorney chooses a_2 and expects a gain in utility of \$10,140, or finally, nature may choose Z_3 as the outcome with probability $P(Z_3) = .095$ and the attorney now chooses a_1 and expects a gain in utility of \$20,832. Therefore, the expected gain in utility to the attorney for choosing the investigation is:

$$(.545)(\$6010) + (.360)(\$10140) + (.095)(\$20832) = \$8,904.89$$

However, from this amount we must subtract the cost of the investigation. This brings up the sticky question of how much are we willing to pay the insurance investigator for carrying out his investigation? Obviously, we will not pay so much that when this amount subtracted from \$8,904.89 leaves an amount less than the \$8,000. we expect to receive without the investigation.

Therefore, the maximum the attorney should be willing to pay for an investigation of this type is \$904.89. If the bill is any higher it would not pay to have this investigation.

CHAPTER V
CONCLUSIONS AND FURTHER SUGGESTIONS

In Chapter I we looked at the statistical decision theory formalism and we saw that we approached the decision making process by dividing the problem into segments and defining sample spaces from each segment of the decision to be made. For example, one segment of the problem that we isolated in such a way was the alternatives that we could take to reach our predefined goal; i.e., the alternatives that are ultimately the object of our decisions. This was not actually a true sample space because no probability measure was defined on it. This is because we follow the convention that we assign probability measure only to those segments which are ultimately decided by nature.

An example of this latter type of segment of the problem that we isolated was the states of nature. Here we could define a probability measure because the decision maker was not making this decision but only predicting it, and the probability measure is representative of his predictions.

In Chapter II we discussed some of the problems involved in applying the theory of Chapter I to the real life world of the legal decision maker.

Chapter III was a hypothetical law case in which the legal decision maker applied the theory of Chapter I in a systematic method of statistical investigation and application and further calculating the worth of the acquisition of

additional information through experimentation. This systematic application of the theory was the model upon which actual law cases were analyzed and finally one was picked for illustration in Chapter IV. This case was analyzed within the confines of Chapter III and a suggestion for gaining further information was evaluated also within the systematic method of our model.

One conclusion I can draw from the above analysis is that although the model may be good for certain legal decision it cannot be extended to include all the possible decisions in the conduct of a law case. There are too many imponderables and before an entire law case can be decided using this model, better means must be found for categorizing the possible courses of action or the possible states of nature.

Another conclusion that I have reached is that for most legal decisions the states of nature can be best described in a model which has a continuous distribution of states of nature. The solution here is a simple extension from discrete to continuous, where the discrete probability measure is replaced by a density function and the expected utility is an integral. The "goal" of maximizing this utility integral will involve the application of variational calculus.

Thus I feel that with the use of continuous distribution of states of nature, continuous density function, and utilities which are functions of the continuous variable

θ and also the use of variational calculus, the extension may be made from the model of this thesis to more complete analysis of legal decisions.

A further conclusion that I have arrived at is the need for criteria for choosing experiments to be analyzed. Some experiments of course are more valuable to perform than others.

In light of the above conclusion, I would suggest further work in the above-mentioned extension of the present model to include a larger number of legal decisions.

There is a need to investigate the various utilities that may occur in the same law case. These utilities will include money, prestige, etc., and a criteria for correlation of these various utilities are needed if they are pertinent to the same decision in a case of if several decisions are to be cascaded in a case. In other words, there should be some attempt at putting utilities on a common scale.

Finally, I would like to add a personal observation in the conduct of the present investigation. This thesis is intended to apply the theory of making statistical decisions to making legal decisions. If the procedure becomes so complex technically that no attorney would be equipped to apply it, what is its value? In addition, it is the practice now of the legal community, when analyzing the use of an investigator as we did statistically in Chapters III and IV, not to

really question the price. They simply take the bill of the investigator and add it on to the bill of their client with no analysis. I wonder if the development of the present theory will change this practice.

APPENDIX

Appropriate Probability Theory

This appendix is devoted to the presentation of the probability theory which was used in the development of this thesis. In our discussion of statistical decision, we have employed the main role played by the concept of "randomness". If the decision maker knew in advance the exact state of nature, there would be no need to apply the concepts of this thesis.

"Randomness" may arise because of the inherent randomness of the process such as is the case in coin tossing, or randomness may arise in a process so complex that it is beyond our ability to ascertain detail. When a process exhibits randomness we talk about its average qualities instead of its exact qualities. Let us now proceed with the presentation of the mathematical tools called the theory of probability with the following definitions.

Sample Space: is any kind of a collection. Each object of the collection is represented by a sample point in this sample space. The sample space is denoted by Ω .

Event: is a group of all sample points (objects of the collection) in the sample space which share a common identifiable attribute. These events are labeled by capital letters A, B, C.... .

Probability Measure: is an assignment of real numbers to the events defined in the sample space. The probability of an event is denoted by $P(A)$, and $0 \leq P(A) \leq 1$.

The definition of a sample space Ω and events such as A, B, \dots implies the existence of certain other identifiable sets of points we are led to the following definitions.

- a. The complement of A , denoted A^c , is the event containing all points in Ω but not in A .

$$A^c = (\omega : \omega \text{ not in } A)$$

- b. The union of A and B , denoted $A \cup B$, is the event containing all points in either A or B or both.

$$A \cup B = (\omega : \omega \text{ in } A \text{ or } B \text{ or both})$$

- c. The intersection of A and B , denoted AB , is the event containing all points in both A and B

$$AB = (\omega : \omega \text{ in both } A \text{ and } B)$$

- d. The event containing no sample points at all is called the null event, denoted \emptyset .

- e. Two events A and B are called disjoint if they contain no common point, i.e., if $AB = \emptyset$.

Since our objective is to use probability theory to predict the results of real-world random experiments, it is reasonable that similar constraints should be imposed upon

corresponding entities in our mathematical model. We therefore restrict our assignment of probability measure to have the following properties.

- I. For every event A_i , $0 \leq \Pr(A_i) \leq 1$
- II. $\Pr(\Omega) = 1$
- III. If $AB = \emptyset$, $\Pr(A \cup B) = \Pr(A) + \Pr(B)$

We define the conditional probability, $\Pr(A/B)$, of an event A given an event B as

$$\Pr(A/B) = \frac{\Pr(AB)}{\Pr(B)} \quad (1)$$

whenever $\Pr(B) \neq 0$. When $\Pr(A)$ is also non-zero, it follows that

$$\Pr(AB) = \Pr(A/B)\Pr(B) = \Pr(B/A)\Pr(A) \quad (2)$$

Clearly, since the intersection of B with itself is B,

$$\Pr(B/B) = 1 \quad (3)$$

Conditional probabilities serve to narrow consideration to a subspace B of a sample space Ω . This is easily visualized pictorially. It is useful to think of "conditioning" as a means of generating a new probability system from a given one.

The new sample space, say Ω' , is the original event B.

The new events, say A_i' , are the original intersections A_iB .

The new probabilities, $\Pr(A_i')$, are the conditional probabilities $\Pr(A_i/B)$.

We see as follows that the probabilities $\Pr(A_i')$ satisfy the required Properties I-III. Since $\Pr(A_iB) \leq \Pr(B)$, $\Pr(A_i') \leq 1$. Also, $\Pr(\Omega') = 1$ by Equation (3). (Division by $\Pr(B)$ in Equation (1) provides the necessary normalization.) Finally, when $A_iA_j = \emptyset$, Property III follows from the equalities

$$\begin{aligned} \Pr(A_iUA_j/B) &= \frac{\Pr(A_iUA_j)B}{\Pr(B)} = \frac{\Pr(A_iBUA_jB)}{\Pr(B)} \\ &= \frac{\Pr(A_iB) + \Pr(A_jB)}{\Pr(B)} = \Pr(A_i/B) + \Pr(A_j/B) \end{aligned}$$

Since conditional probabilities can be considered as ordinary probabilities on a new sample space, all statements and theorems about ordinary probabilities also hold true for conditional probabilities. In particular, if the set of intersections (A_iB) are a disjoint partitioning of B then

$$\Pr(B) = \sum_{\text{all } i} \Pr(A_iB) = \sum_{\text{all } i} \Pr(B)\Pr(A_i/B) \quad (4)$$

and

$$1 = \sum_{\text{all } i} \Pr(A_i/B) \quad (1) \quad (5)$$

(1) 6.311 Notes by Jacobs & Wozencraft, 1963.

To sum up this appendix on appropriate probability theory let us consider the following application.

I have a bowl with 10 white and 6 black marbles in it. I am going to reach in and draw a second marble at random. The question is what is the probability that I will draw one white marble and one black marble. This can happen in two mutually exclusive ways; i.e., I can draw a white marble on the first draw and a black marble on the second draw, or I can draw a black marble on the first draw and a white marble on the second draw. Therefore:

$$P(\text{drawing one white and one black marble}) =$$

$$P(\{\text{white on 1st draw}\} \{\text{black on 2nd draw}\}) +$$

$$P(\{\text{black on 1st draw}\} \{\text{white on 2nd draw}\})$$

By equation (2) above, we have:

$$P(\{\text{white on 1st draw}\} \{\text{black on 2nd draw}\}) =$$

$$P(\{\text{black on 2nd draw}\} / \{\text{white on 1st draw}\}) P(\text{white on 1st draw})$$

$$= \left(\frac{6}{15}\right) \left(\frac{10}{16}\right) = .25$$

also

$$P(\{\text{black on 1st draw}\} \{\text{white on 2nd draw}\}) =$$

$$P(\{\text{white on 2nd draw}\} / \{\text{black on 1st draw}\}) P(\text{black on 1st draw})$$

$$= \left(\frac{10}{15}\right) \left(\frac{6}{16}\right) = .25$$

therefore,

$$P(\text{drawing one white and one black marble}) =$$
$$(.25) + (.25) = .50$$

BIBLIOGRAPHY

- (1) Chernoff, Herman and Lincoln E. Moses; Elementary Decision Theory, New York, Wiley (1959)
- (2) Drescher, Melvin; Games of Strategy: Theory and Applications, Englewood Cliffs, New Jersey, Prentice Hall (1961)
- (3) Feller, William; An Introduction to Probability Theory and its Applications, New York, Wiley (1957 -
- (4) Luce, R. Duncan and Howard Raiffa; Games and Decisions: Introduction and Critical Survey, New York, Wiley (1957)
- (5) Machal, Robert E. ed; Information and Decision Processes, New York, McGraw-Hill (1960)
- (6) Parzen, Emanuel; Modern Probability Theory and Its Applications, New York, Wiley (1960)
- (7) Raiffa, Howard and Robert Schlaifa; Applied Statistical Decision Theory, Harvard University (1961)
- (8) Williams, John Davis; The Complete Strategyst, New York, McGraw-Hill, (1954)

Other References Used By the Author

- (1) M.U.L.L., Modern Uses of Logic In Law; is the newsletter of the American Bar Association Special Committee on Electronic Data Retrieval, and is published in collaboration with Yale Law School.
- (2) Statistical Decision Theory Class Notes; (by) Prof. Jacobs and Prof. Wozencraft of M.I.T.