QUANTITATIVE METHODS

Decision Tree

Rabies Example

Background Information

Rabies, a disease of the central nervous system, is generally spread by domestic dogs, bats, and wild animals. The symptoms of the disease normally appear 4 to 6 weeks after exposure. The patient is irritable, may have convulsions and finally goes into a coma. Death often occurs 3 to 5 days after the symptoms appear.

In 1881, Louis Pasteur found that the infectious agent could be recovered from the brain of a dead rabies-infected animal. The agent was invisible under the microscopes then available and could not be cultivated in nutrients. The agent was called a virus (Latin for poison). After much experimentation Pasteur developed a vaccine that he hoped would prevent the disease. In 1885, Pasteur successfully treated a peasant boy who was bitten by a rabid dog.

The treatment of rabies today consists of one injection of antirabies globulin followed by five injections of rabies vaccine.

Case Details

Our subject has been bitten by a small dog during an evening in Tijuana, Mexico. The dog escaped and there is no hope of finding it. Should our subject undergo the rabies treatment?

(This is based on the autobiographical account by Freeman Dyson.)

Data

1. 18% of bites by a rabid dog cause rabies if untreated.

2. Anyone who contracts rabies dies.

3. Of all people given the series of rabies injections,

   1 in 200 contract rabies before the vaccine can take effect,
   1 in 600 contract encephalitis from the injections and death ensues.

4. We need to know the probability that the dog is rabid. That is, what fraction of the dogs which bite pedestrians in Tijuana are rabid? We can make an assumption that this is 1/2, but we will want to see how the final decision depends on this probability (that is, what is this particular sensitivity?).

Structure of the Decision Tree

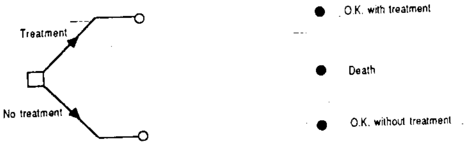
There is only one immediate decision to be made -- whether to have the treatment or not.

There are three possible outcomes of our decision:

1. Our subject is O.K. after the treatment.

2. Our subject dies.

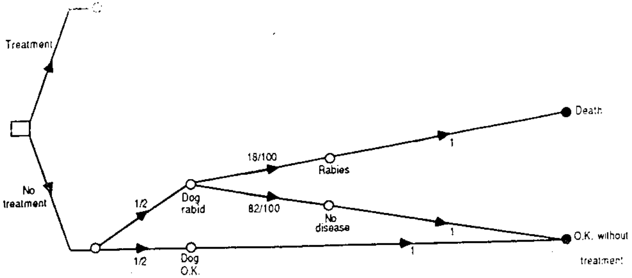
3. Our subject is O.K. without the treatment.



To construct the decision tree, we must now fill in the various ways in which we can move from either decision to one of the outcomes.

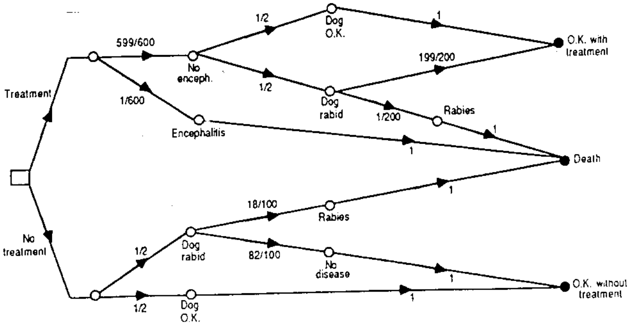
Partial Tree after no Treatment

The easier half of the decision tree is that connecting the "No treatment" decision to the outcomes. Once we decide on no treatment, we know half the cases will be O.K. since the dog is not rabid. Furthermore, even if the dog is rabid, 82% of the cases will be O.K. Thus, we can draw this portion of the decision tree.



Remainder of Tree

Now we turn to the top half of the decision tree -- the events that follow a decision to have treatment. This portion is shown below. Here we assume that we first can separate off the group that contracts encephalitis.



For example, in the bottom half of the figure, if the dog is rabid, there is a probability of 82/100 our subject will be O.K. without treatment. This branch at this node has a value

$$\underbrace{\frac{82}{100}}_{\text{Probability}} \times \underbrace{100}_{\text{Value of outcome}} \quad \text{or} \quad 82$$

The other branch has a value

$$\frac{18}{100} \times 0 \quad \text{or} \quad 0$$

Thus, the node labelled "Dog rabid" has the value

$$82 + 0 \quad \text{or} \quad 82$$

In this way we fill in the numbers shown in the diagram and find that the expected value of the "No treatment" option is 91.

Similarly, for the top half of the diagram we calculate *backwards* again to obtain the numbers shown on the figure, and finally the value of the "Treatment" option as 79.7.

Thus, the decision tree shows that the preferable choice is the "No treatment" option.

Sensitivity Study

Two sensitivity studies are important in this case:

How does the optimum decision depend on the value we chose (80) for the "O.K. with treatment" outcome? If we calculate backward, we find that a value greater than 91.4 for the "O.K. with treatment" outcome results in a preference for the treatment option. Obvioŭsly, the problem is quite sensitive to this value which we are estimating.

In the section on Data we assumed that 1/2 of the dogs which bite pedestrians in Tijuana are rabid. If the actual fraction is not 1/2, the optimum decision could be different.

Purpose of Decision Trees

The real purpose of decision analysis is to force us to think of all the factors that are important in making a decision. The decision tree requires that we look at the process step-by-step to see which probabilities are critical and to spotlight the values associated with different outcomes.

		Election victory: .71 Conservative	.26 Labor
Action on stock	Sell	100	100
	Hold	300	−200
	Buy	600	−400

$$E(\text{Sell}) = .71(100) + .29(100) = 100$$

$$E(\text{Hold}) = .71(300) + .29(-200) = \frac{5900}{38} = 155.26$$

$$E(\text{Buy}) = .71(600) + .29(-400) = 310.53$$

Therefore, in order to maximize expected profit, Mrs. Magoo should buy more stock. Note that this is not the same decision as that based on the original probabilities. In the solution in Section 15.3, sell had the highest expected value.

Bayesian decision analysis consists essentially of revising the prior (objective or subjective) probabilities concerning the states of nature in view of additional data or information and basing the decision on these revised probabilities. In the preceding example the conditional probabilities of the reporter's predictions given that certain election outcomes were going to occur were of a rather subjective nature. If the reporter had made predictions for a large number of elections, these conditional probabilities could, of course, have been based on (objective) relative frequencies. In many problems relatively objective determination of these probabilities is possible, as in the following example.

EXAMPLE
The Blue Bolt Company has had a series of mix-ups on its production line, and as a result there is a case containing 100,000 screws in the shipping room ready for shipment but unmarked as to size, production run, or other identification. The company produces three sizes of screws having lengths whose means and standard deviations are, respectively, 5 and 1.5, 6 and 2.5, and 8 and 4 in. The company must decide what to do with the case of screws; there are costs involved in sending the case if it is not the size ordered and costs involved in discarding the case. These costs (including loss of good will and partial replacements costs when relevant) are given by the following matrix:

	$\frac{1}{3}$ $\mu = 5$	$\frac{1}{3}$ $\mu = 6$	$\frac{1}{3}$ $\mu = 8$	Expected value
Ship as $\mu = 5$	0	−300	−500	−800/3 ←
Ship as $\mu = 6$	−1600	0	−400	−2000/3
Ship as $\mu = 8$	−2000	−1800	0	−3800/3
Discard	−1000	−1200	−1500	−3700/3

(a) Assuming that the screws are equally likely to be each of the three sizes, what should the Blue Bolt Company do to maximize its expected value? (b) Assuming that the lengths of each size of bolts are normally distributed what should the company do if a random sample of 100 bolts has a mean length of 6.8 in.? (Consider 6.8 as including the interval 6.8 ± .1 to solve the problem of continuity;—this can be justified on the basis of errors of measurement).

(a) The company would ship the screws to fill an order for screws having an average length of 5 in., since this decision has the minimum expected loss (given at the side of the above table).

(b) Given the sample of size 100 with $\bar{x} = 6.8$ in., the probability of this sample coming from the three possible distributions would be calculated as follows

$$n = 100$$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{6.7 - 5}{.15} = \frac{1.7}{.15} = 12.6;\ P(\bar{x} = 6.8 \pm .1 \mid \mu = 5, \sigma = 1.5) = 0$$

$$\left.\begin{aligned} z &= \frac{6.9 - 6}{.25} = \frac{.9}{.25} = 3.6 \\ z &= \frac{6.7 - 6}{.25} = \frac{.7}{.25} = 2.8 \end{aligned}\right\} P(2.8 \le z \le 3.6 \mid \mu = 6, \sigma = 2.5) = .0026 - .0011 = .0015$$

$$\left.\begin{aligned} z &= \frac{6.9 - 8}{.4} = \frac{-1.1}{.4} = -2.75 \\ z &= \frac{6.7 - 8}{.4} = \frac{-0.9}{.4} = -2.25 \end{aligned}\right\} P(-2.75 \le z \le -2.25 \mid \mu = 8, \sigma = 4) = 0.122 - .0030 = .0092$$

$$P(5 \mid 6.8) = \frac{P(6.8 \mid 5)P(5)}{P(6.8)} = \frac{0(\frac{1}{3})}{0 + (.0015)(\frac{1}{3}) + (.0092)(\frac{1}{3})} = 0$$

$$P(6 \mid 6.8) = \frac{P(6.8 \mid 6)P(6)}{P(6.8)} = \frac{.0005}{.00357} = .14$$

$$P(8 \mid 6.8) = \frac{P(6.8 \mid 8)P(8)}{P(6.8)} = \frac{.00307}{.00357} = .86$$

The following table summarizes the prior and posterior probability distributions for this example:

EVENT (X)	PRIOR PROBABILITY $P(X)$	CONDITIONAL PROBABILITY OF SAMPLE $P(S \mid X)$	JOINT PROBABILITY OF SAMPLE AND EVENT $P(S,X)$	POSTERIOR PROBABILITY $P(X \mid S)$
$\mu = 5$, $\sigma = 1.5$	1/3	0	0	0
$\mu = 6$, $\sigma = 2.5$	1/3	.0015	.0005	.14
$\mu = 8$, $\sigma = 4$	1/3	.0092	.00307	.86
			.00357	1.00

Considering the decision payoff matrix in view of these revised (posterior) probabilities, we have

	0	.14	.86	
	$\mu = 6$	$\mu = 6$	$\mu = 8$	Expected value
Ship as $\mu = 5$	0	−300	−500	−472
Ship as $\mu = 6$	−1600	0	−400	−344
Ship as $\mu = 8$	−2000	−1800	0	−252
Discard	−1000	−1200	−1500	−1458

So the company would ship the screws to fill an order for screws having an average length of 8 in.

16.3 TESTING SCIENTIFIC AND STATISTICAL HYPOTHESES

A statistical hypothesis is a statement about the distribution of a random variable; most statistical hypotheses concern the value of a parameter of a distribution of known or assumed form, although some hypotheses concern the form of the distribution itself. A test of a statistical hypothesis is a procedure for deciding whether to accept or reject the hypothesis on the basis of the outcome of a random experiment—that is, on the basis of the observed value of a random variable.

The use of a random experiment to test a statistical hypothesis is based on an extension of the method of testing scientific hypotheses by nonrandom experiments. For centuries, hypotheses in the physical sciences have been tested by experiments. Galileo reportedly dropped a large and a small cannon ball from the leaning tower of Pisa in order to test the hypothesis that two objects of different weights fall at the same speed. His hypothesis was supported by the fact that the cannon balls did in fact strike the ground at nearly the same instant. More recently and more spectacularly, Einstein's hypothesis that energy equals mass times the square of the speed of light was supported by the explosion of the atomic bomb.

The logic of testing scientific hypotheses by nonrandom experiments is as follows: According to the hypothesis, a specific result should occur or an observable quantity should have a specified value. If the occurrence predicted by the hypothesis is observed when the experiment is performed, then the hypothesis is supported. It is not proved, however, since there may be other hypotheses that make the same prediction. This procedure is based on the assumption that the experiment is performed precisely as required to test the hypothesis and that there is no error in measurement or observation—that is, it is based on the assumption that there is a "critical experiment" for the hypothesis.

Even under ideal circumstances with regard to conduct of experiments and precision of measurement, a theory or hypothesis can never be proved. As data consistent with a theory or hypothesis accumulate and contradictory data are not observed, the theory or hypothesis is supported and may even eventually be regarded as proved for practical purposes. However, one counterexample or one ond, the calculation of $E(D_1)$, $E(D_2)$, and $E(D_3)$ did not require conducting ıy physical experiment. This procedure is frequently called *preposterior*: the process lows conceptual treatment of the problem before you have to make a decision.

Third, the cost of sample information (or the cost of testing a switch) was given. he payoffs associated with d_1, d_2, d_3, and d_4 incorporated the $5-per-switch cost of sting. Frequently, decision problems will not have that information, and the iestion is: What is the maximum amount the decision maker should pay for sample mperfect) information?

ı.6 EXPECTED VALUE OF SAMPLE INFORMATION

n example will serve to demonstrate the basic process for evaluating the expected ılue of sample information.

Example 14.2 A Souvenir Tennis Program

Ms. Robbie Biggs has the exclusive rights to produce "the souvenir pamphlet" for a challenge tennis match. Two months before the match Ms. Biggs had 5,000 copies of the souvenir program printed at a cost of $5,200.

The program includes the rules for the match, score sheet, and some tennis history, but the program is primarily a biography of the players. The $2 per copy sale price was boldly printed on the cover. Three weeks before the match, a general sports magazine prints a biographical sketch of the players that essentially duplicates the material in the souvenir program. Ms. Biggs is discouraged and reassesses the probability associated with the levels of sales for the program.

Ms. Alacrity offers to pay Ms. Biggs $6,200 for the souvenir programs and the ˌusive rights to distribute them. Mr. Fondant offers to pay Ms. Biggs $6,100 and 20 percent of any profit. All agree that the program will sell for $2 per copy. Ms. Biggs is faced with the following decisions:

D_1: Sell rights and printed copies to Ms. Alacrity
D_2: Sell rights and printed copies to Mr. Fondant
D_3: Keep rights and sell them herself

TABLE 14.10 PAYOFF TABLE ($ PROFIT) FOR MS. BIGGS

Ms. Biggs' Probability Assessment $P(SON_i)$	Number of Copies Demanded SON_i	Decisions: Sell to Alacrity D_1	Sell to Fondant D_2	Sell Them Herself D_3
0.50	2,000	1,000	900	−1,200
0.20	3,000	1,000	900	800
0.20	4,000	1,000	1,280	2,800
0.10	5,000	1,000	1,680	4,800

Ms. Biggs considers only four possible states of nature: 2,000 copies, 3,000 copies, 4,000 copies, or 5,000 copies will be sold. If she takes Ms. Alacrity's offer, Ms. Biggs will receive \$6,200 − \$5,200 = \$1,000 profit regardless of the number of souvenir programs sold.

If Ms. Biggs sells the programs herself, her profit function is:

$$\text{Profit} = (\$2 \text{ per copy})(\text{number of copies sold}) - \$5{,}200$$

For the four states of nature, the profits are shown under D_3 in Table 14.10.

For 2,000 copies sold, there is a loss. If Mr. Fondant's offer is accepted, Ms. Biggs will receive \$6,100 − \$5,200 = \$900 profit for that state of nature. If Mr. Fondant makes a profit, then Ms. Biggs will receive

$$\$900 + (0.20)[(\$2)(\text{number of copies sold}) - \$6{,}100]$$

or \$900 plus 20 percent of Mr. Fondant's profit. If only 3,000 copies are sold, then Mr. Fondant's revenue is \$6,000, but his costs were \$6,100. If 4,000 copies are sold, then Mr. Fondant's profit is \$8,000 − \$6,100 = \$1,900, of which 20 percent is \$380; thus, Ms. Biggs would receive \$900 + \$380 = \$1,280 in profit.

From Table 14.10, the minimax decision is D_1 and the maximax decision is D_3. From the expected-value calculations in Table 14.11, the optimal expected value is \$1,054 (recall that optimal is "maximum" in this problem because the payoffs are "profits").

TABLE 14.11 EXPECTED VALUE CALCULATIONS

$P(SON_i)$	D_1	D_2	D_3	(1) × (2)	(1) × (3)	(1) × (4)	$P(SON_i) \cdot (V_{max} \mid SON_i)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0.50	1,000	900	−1,200	500	450	−600	500
0.20	1,000	900	800	200	180	160	200
0.20	1,000	1,280	2,800	200	256	560	560
0.10	1,000	1,680	4,800	100	168	480	480
				1,000	1,054	600	1,740
				$E(D_1)$	$E(D_2)$	$E(D_3)$	$E(CP)$

Since the expected value of certain prediction is \$1,740, then the expected value of perfect information is

$$E(PI) = 1{,}740 - 1{,}054 = \$686$$

Now, let us introduce another wrinkle in Ms. Biggs' decision problem. Most of the ticket sales for the match were transacted by mail. A marketing research firm offers to randomly select a sample of n persons from this population, find out how many of those n people will buy a souvenir program, and report the results to Ms. Biggs. How much should she pay for that sample information? Certainly not more than \$686, the expected value of perfect information. Perhaps we can obtain a better evaluation of the worth of imperfect information.

First, let us make the following assumptions:

1. There will be 10,000 people attending the tennis match.
2. The proportion of attendees who will buy a souvenir program is represented by π.
3. The proportion of mail-order ticket recipients who will buy a souvenir program is also equal to π.
4. Ms. Biggs' assessment of the states of nature is correct. (That is, exactly 2,000, 3,000, 4,000, or 5,000 programs will be sold.)
5. A sample of size ($n =$) 4 will be selected from the mail-order ticket recipients.
6. Although the sample will be selected without replacement, the sample size is so small relative to the population size that binomial probabilities can be used to calculate the probability of the various sample results.

Second, let us define D_4 as the "sample of size 4 plan." That is, the marketing research firm will select a random sample of size 4 and report to Ms. Biggs that either zero, one, two, three, or four of those people will buy a souvenir program. Based upon what they report, Ms. Biggs will choose the decision (i.e., sell rights to Ms. Alacrity, sell to Mr. Fondant, or keep the rights) with the highest expected value. To keep track of all this, consider the decision tree in Figure 14.4.

Finally, let ${}_kd_j$ = decision D_j given the sample result that k people will buy the souvenir program. Then, the expected value for ${}_kd_j$ is

$$E({}_kd_j) = \sum_{i=1}^{4} [V_i \cdot P(\text{SON}_i \mid k)]$$

where $P(\text{SON}_i \mid k)$ is the posterior probability. For example, the probability that 5,000 copies will be sold (SON_4) given that ($k = 2$) two of the four people sampled will buy the souvenir program is $P(\text{SON}_4 \mid k = 2)$. To obtain the posterior probabilities (via Bayes' rule), we need the conditional probabilities: $P(k \mid \text{SON}_i)$. Using the assumptions stated above, we can use the binomial probabilities from Appendix Table III for the conditional probabilities.

How? If 2,000 people will buy the program, which represents SON_1, then $\pi = 2{,}000/10{,}000 = 0.20$. Since $n = 4$,

$$P(k \mid \text{SON}_1) = P(k \mid \pi = 0.20, n = 4)$$

In general, $P(k \mid \text{SON}_i) = P(k \mid \pi = \text{SON}_i/10{,}000,\ n = 4)$. Table 14.12 presents the appropriate binomial probabilities from Appendix Table III.

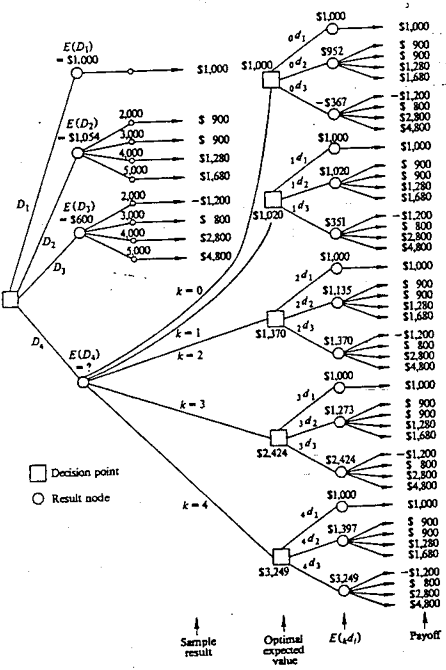


FIGURE 14.4 Decision tree for Ms. Biggs.

TABLE 14.12 CO

Number of People in Sample Who Will Buy k
0
1
2
3
4

Consider the sar
will buy a souvenir p
tion where 5,000 (or
find that

$P($

For $k = 0$, we need tc
assessment of which s
and Bayes' rule to ot

Table 14.13 serves as
(2) calculating the exp
rights. Regardless of t
the decision to sell to

Finally, we are at
4 plan." For a given s
with the optimal e
$951.78 > E({}_0d_3) = -$
of getting each sample
shown in Table 14.14.

The expected val
expected values for th
pling cost? (How muc

ond, the calculation of $E(D_1)$, $E(D_2)$, and $E(D_3)$ did not require conducting ny physical experiment. This procedure is frequently called *preposterior*: the process lows conceptual treatment of the problem before you have to make a decision.

Third, the cost of sample information (or the cost of testing a switch) was given. he payoffs associated with d_1, d_2, d_3, and d_4 incorporated the $5-per-switch cost of sting. Frequently, decision problems will not have that information, and the uestion is: What is the maximum amount the decision maker should pay for sample mperfect) information?

4.6 EXPECTED VALUE OF SAMPLE INFORMATION

n example will serve to demonstrate the basic process for evaluating the expected ilue of sample information.

Example 14.2 A Souvenir Tennis Program

Ms. Robbie Biggs has the exclusive rights to produce "the souvenir pamphlet" for a challenge tennis match. Two months before the match Ms. Biggs had 5,000 copies of the souvenir program printed at a cost of $5,200.

The program includes the rules for the match, score sheet, and some tennis history, but the program is primarily a biography of the players. The $2 per copy sale price was boldly printed on the cover. Three weeks before the match, a general sports magazine prints a biographical sketch of the players that essentially duplicates the material in the souvenir program. Ms. Biggs is discouraged and reassesses the probability associated with the levels of sales for the program.

Ms. Alacrity offers to pay Ms. Biggs $6,200 for the souvenir programs and the usive rights to distribute them. Mr. Fondant offers to pay Ms. Biggs $6,100 and 20 percent of any profit. All agree that the program will sell for $2 per copy. Ms. Biggs is faced with the following decisions:

D_1: Sell rights and printed copies to Ms. Alacrity
D_2: Sell rights and printed copies to Mr. Fondant
D_3: Keep rights and sell them herself

TABLE 14.10 PAYOFF TABLE ($ PROFIT) FOR MS. BIGGS

		Decisions		
Ms. Biggs' Probability Assessment $P(SON_i)$	Number of Copies Demanded SON_i	Sell to Alacrity D_1	Sell to Fondant D_2	Sell Them Herself D_3
0.50	2,000	1,000	900	−1,200
0.20	3,000	1,000	900	800
0.20	4,000	1,000	1,280	2,800
0.10	5,000	1,000	1,680	4,800

TABLE 14.13 POSTERIOR PROBABILITY AND EXPECTED-VALUE CALCULATIONS

A Priori Probability $P(SON_j)$ Table 14.10 (1)	Conditional Probability $P(k\|SON_j)$ Table 14.12 (2)	Joint Probability $P(k \cap SON_j)$ (1) × (2) (3)	Posterior Probability $P(SON_j\|k)$ (3) ÷ $P(k)$ (4)	Payoff from D_2 Table 14.10 (5)	$E(_kd_j)$ Calculation (4) × (5) (6)	Payoff from D_3 Table 14.10 (7)	$E(_kd_j)$ Calculation (4) × (7) (8)
$k = 0$							
0.50	0.4096	0.2048	0.7186	900	646.74	−1,200	−862.32
0.20	0.2401	0.0480	0.1684	900	151.56	800	134.72
0.20	0.1296	0.0259	0.0909	1,280	116.35	2,800	254.52
0.10	0.0625	0.0063	0.0221	1,680	37.13	4,800	106.08
	$P(k=0)$ =	0.2850	1.0000	$E(_0d_2)$ =	$951.78	$E(_0d_3)$ =	−367.00
$k = 1$							
0.50	0.4096	0.2048	0.5372	900	483.48	−1,[illegible]00	−644.64
0.20	0.4116	0.0823	0.2159	900	194.31	[illegible]00	172.72
0.20	0.3456	0.0691	0.1813	1,280	232.06	2,800	507.64
0.10	0.2500	0.0250	0.0656	1,680	110.21	4,800	314.88
	$P(k=1)$ =	0.3812	1.0000	$E(_1d_2)$ =	$1,020.06	$E(_1d_3)$ =	350.60
$k = 2$							
0.50	0.1536	0.0768	0.3250	900	292.50	−1,200	−390.00
0.20	0.2646	0.0529	0.2239	900	201.51	800	179.12
0.20	0.3456	0.0691	0.2924	1,280	374.27	2,800	818.72
0.10	0.3750	0.0375	0.1587	1,680	266.62	4,800	761.76
	$P(k=2)$ =	0.2363	1.0000	$E(_2d_2)$ =	$1,134.90	$E(_2d_3)$ =	1,369.60
$k = 3$							
0.50	0.0256	0.0128	0.1531	900	137.79	−1,200	−183.72
0.20	0.0756	0.0151	0.1806	900	162.54	800	144.48
0.20	0.1536	0.0307	0.3672	1,280	470.02	2,800	1,028.16
0.10	0.2500	0.0250	0.2990	1,680	502.32	4,800	1,435.20
	$P(k=3)$ =	0.0836	0.9999	$E(_3d_2)$ =	$1,272.67	$E(_3d_3)$ =	2,424.12
$k = 4$							
0.50	0.0016	0.0008	0.0580	900	52.20	−1,200	−69.60
0.20	0.0081	0.0016	0.1159	900	104.31	800	92.72
0.20	0.0256	0.0051	0.3696	1,280	473.09	2,800	1,034.88
0.10	0.0625	0.0063	0.4565	1,680	766.92	4,800	2,191.20
	$P(k=4)$ =	0.0138	1.0000	$E(_4d_2)$ =	$1,396.52	$E(_4d_3)$ =	3,249.20

TABLE 14.14 EXP
SAMPLING COSTS)

Sample Result k (1)
0
1
2
3
4

Without sampli
selling to Mr. Fonda
to pay for the sampl

(Max
of samp

If the cost of sa
criterion would sugg
firm to obtain the sa

$k = 0$

$k = 1$

$k = 2,$

If the cost of s
expected value criter
Note that this a
$n = 5$, we would star
14.12. Then for each
recalculated. As in
sample result and P(

TABLE 14.14 EXPECTED VALUE OF SAMPLE PLAN (EXCLUDING SAMPLING COSTS)

Sample Result k (1)	Probability of Sample Result (Table 14.13) $P(k)$ (2)	Optimal Expected Value Given Sample Result (Figure 14.4) $[E(\cdot, d_j)]_{max}$ (3)	Expected Value Calculation (2) × (3) (4)
0	0.2850	\$1,000.00 (d_1)	\$ 285.00
1	0.3812	1,020.06 (d_2)	388.85
2	0.2363	1,369.60 (d_3)	323.64
3	0.0836	2,424.12 (d_3)	202.66
4	0.0138	3,249.20 (d_3)	44.84
	0.9999		\$1,244.99 = $E(D_4)$

Without sampling information, the optimal expected value is associated with selling to Mr. Fondant, $E(D_2)$ = \$1,054. Thus, the most Ms. Biggs would be willing to pay for the sample information is

$$\begin{pmatrix}\text{Maximum value}\\ \text{of sample information}\end{pmatrix} = \begin{pmatrix}\text{optimal expected value}\\ \textit{with}\text{ sample information}\end{pmatrix} - \begin{pmatrix}\text{optimal expected value}\\ \textit{without}\text{ sample information}\end{pmatrix}$$

$$= \$1{,}244.99 - \$1{,}054.00 = \$190.99$$

If the cost of sampling is less than \$190.99, then the optimal expected value criterion would suggest the following decision rule. Engage the marketing research firm to obtain the sample information. If they report

$k = 0$	then sell rights to Ms. Alacrity
$k = 1$	then sell rights to Mr. Fondant
$k = 2$, 3, or 4	then Ms. Biggs should keep the rights

If the cost of sampling is more than \$190.99 for a sample of 4, the optimal expected value criterion would suggest that Ms. Biggs sell the rights to Mr. Fondant.

Note that this analysis is only specific for a sample of size 4. Consider $n = 5$. For $n = 5$, we would start with the binomial probabilities for $n = 5$ and reconstruct Table 14.12. Then for each sample result (that is, $k = 0, 1, \ldots, 5$), Table 14.13 would be recalculated. As in Table 14.14, we would use the optimal expected value for each sample result and $P(k)$ to find $E(D_5)$.

In general, the expected value for "a plan using a sample of size n" will vary with the size of the sample, or $E(D_n) = f(n)$. Also, the cost of sampling usually varies with the size of sample, or $C_n = g(n)$. Let $E(D_j^*)$ represent the optimal expected value of those decisions which do not use sampling information. Then the *net expected gain from sampling* (NEGS) is

$$\text{NEGS} = E(D_n) - C_n - E(D_j^*)$$

which varies with n, the size of the sample.

If, for all values of n, NEGS is less than 0, then the optimal expected value criterion yields D_j^* as the optimal decision (i.e., do not sample).

If NEGS is greater than 0 for one or more values of n, then the optimal expected value criterion suggests that you should use sample information and use a sample size that provides the largest value of NEGS. Hence, the procedures discussed above not only assess the expected value of sample information but also provide a basis for deciding what size sample should be employed.

SUMMARY

This concludes the introduction to Bayesian decision analysis. The essential power of this logic, or technique, should be obvious. Most of the material in the preceding chapters can be classified as "classical statistics." The techniques in this chapter share some of that material, but there are differences which have fired controversy. That controversy has produced both heat and light. For those who progress beyond the scope of this book, do not ignore the light of Bayesian analysis.

In retrospect, the logic of (1) identifying what could happen (the states of nature), (2) deciding what alternatives the decision maker has (the decisions), and (3) measuring the payoff for each decision, given a state of nature, in units that are most appropriate to the goals of the decision maker, seems almost trivial. But implementing that logic forces the decision maker to be very explicit about the problem. And, as we have seen, the payoff table is only the beginning.

Assigning a value to the probability of occurrence for each state of nature may require the use of subjective probabilities. There are techniques (not covered here) for checking the decision maker's "degree of belief."

Bayes' rule is a device for revising those initial probability assessments in the light of sample information. Finally, the expected value of information (either perfect or imperfect) provides a conceptual mechanism for evaluating sample information *before* that information is obtained.

Remember, this is only an introduction—an overview and an enticement. For those who seek it, there is more.

Values for Outcomes

In order to compare quantitatively the two options in the decision, we need to assign values to each of the three outcomes. Notice that the decision tree above already serves an important purpose even if we stop before assigning values: the tree shows clearly all the various factors and data which should be included as our subject tries to make a decision. Thus, the decision tree helps to organize our thinking about a problem.

In decision-tree problems, it is common (but not necessary) to assign values for the various outcomes in the range 0-100. If we follow this practice,

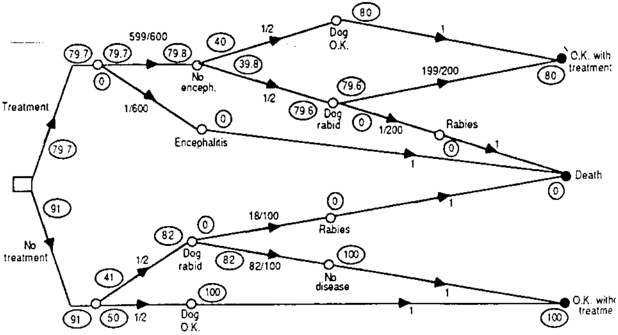
"Death" obviously has the value 0

"O.K. without treatment" has the value 100.

The question is what value to give "O.K. with treatment" -- this is obviously less than 100 (the treatment is painful and costly), but far more than 0.

There are various techniques to decide on this unknown value, but basically it is the question of personal judgment. To continue the analysis, we select the value 80. (Later we may want to study how sensitive the decision is to this particular value.)

We now have the complete decision tree, shown below.



Values shown encircled

Finding the Optimum Decision

To find the optimum decision, we determine the *expected value* for each of the two options by working backward from the outcomes toward the decision point.