# 統計學

Spring 2026

政治大學統計系余清祥

2026年05月19日

第十九章：決策分析

http://csyue.nccu.edu.tw

# Chapter Contents



CENGAGE

# Introduction

**Decision analysis** can be used to develop an optimal decision strategy when faced with several decision alternatives and an uncertain or risk-filled pattern of future events.

We begin the study of decision analysis by considering decision problems that involve reasonably few decision alternatives and reasonably few future events.

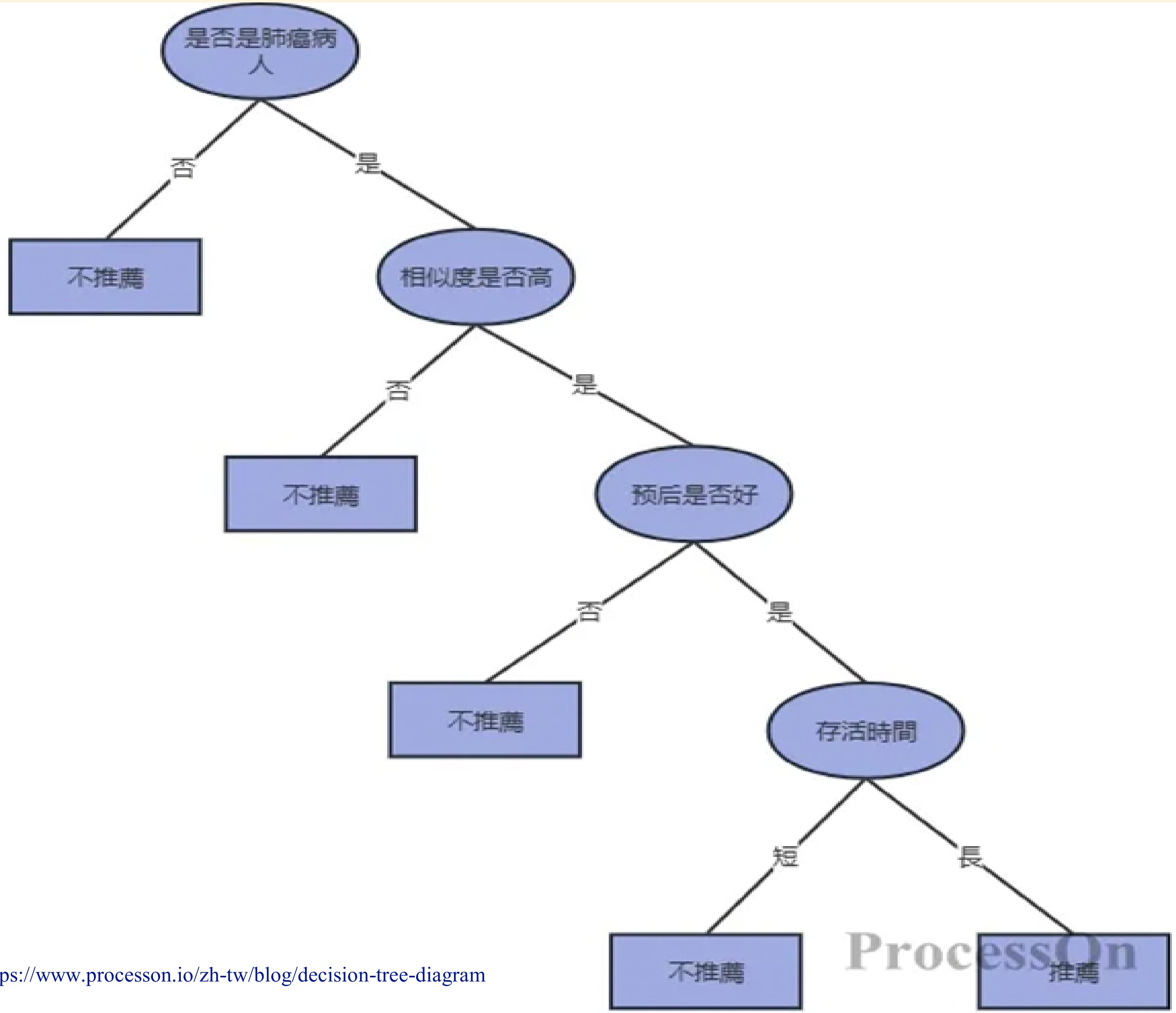
**Payoff tables** are introduced to provide a structure for decision problems.

We then introduce **decision trees** to show the sequential nature of the problems.

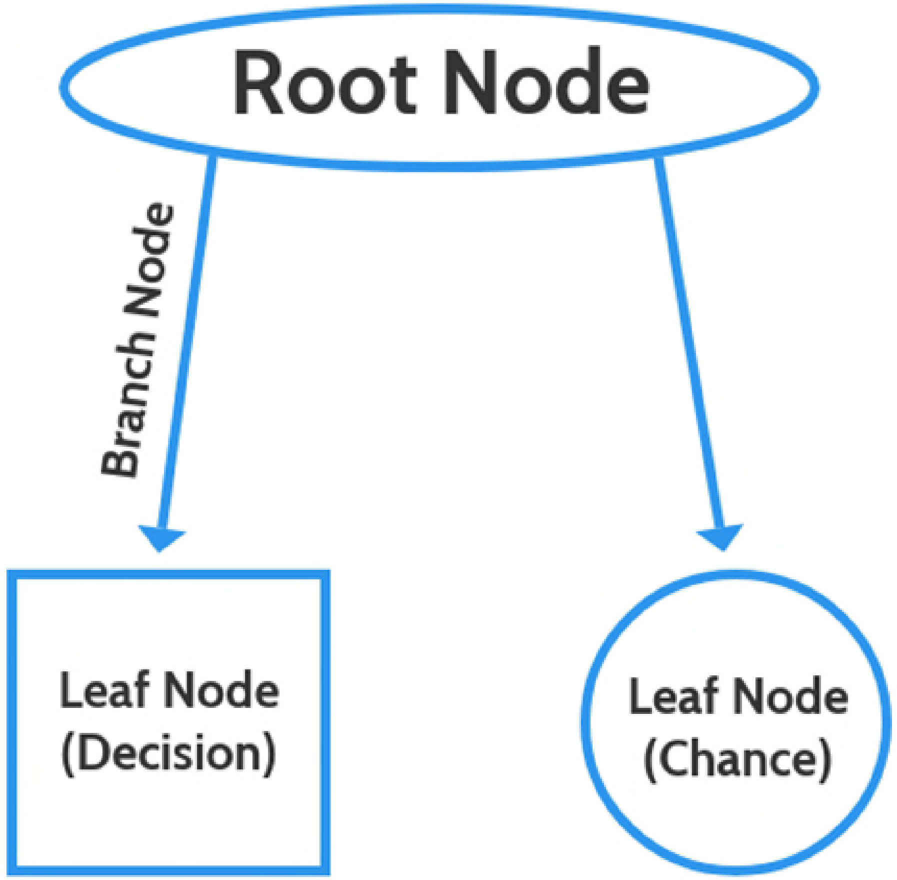
Decision trees are used to analyze more complex problems and to identify an optimal sequence of decisions, referred to as an optimal decision strategy.

In the last section, we show how **Bayes' theorem** can be used to compute branch probabilities for decision trees.

# 什麼是決策樹？

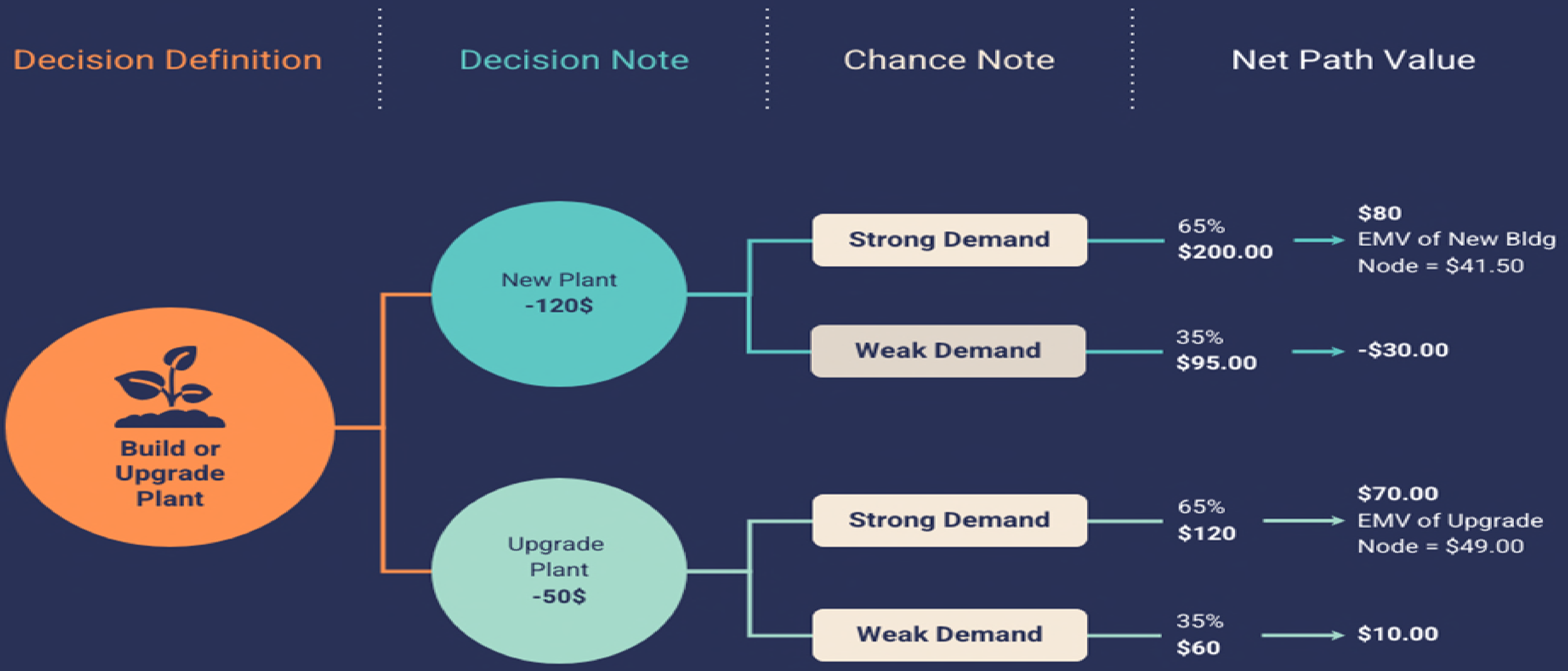



https://www.processon.io/zh-tw/blog/decision-tree-diagram




https://venngage.com/blog/what-is-a-decision-tree/

# Decision Tree Analysis




https://venngage.com/blog/what-is-a-decision-tree/

# 19.1 Problem Formulation

Problem formulation, the first step in the decision analysis process, begins with a verbal statement of the problem.

A decision problem is characterized by decision alternatives, states of nature, and payoffs:

- The **decision alternatives** are the possible strategies the decision maker can employ.
- The uncertain future events are referred to as *chance events*. The one and only outcome associated to each chance event is called a **state of nature**.
- The **payoff** is the consequence resulting from a specific combination of a decision alternative and a state of nature.

Let us begin our study of the decision analysis process by considering a construction project for the Pittsburgh Development Corporation (PDC.)

PDC purchased land to build a new luxury condominium complex. PDC is considering three different projects with either 30, 60, or 90 condominiums.

# 19.1 Problem Statement

The financial success of the PDC project depends upon the size of the condominium complex and the chance event concerning the demand for the condominiums.

PDC problem statement is to select the size of the new luxury condominium project that will lead to the largest profit given the uncertainty concerning the demand for the condominiums.

Given the problem statement, PDC has the following three decision alternatives:

$d_1$ = a small complex with 30 condominiums

$d_2$ = a medium complex with 60 condominiums

$d_3$ = a large complex with 90 condominiums

The chance event concerning the demand for the condominiums has two states of nature:

$s_1$ = strong demand for the condominiums

$s_2$ = weak demand for the condominiums

# 19.1 Payoff Table

PDC management must first select a decision alternative (complex size), then a state of nature follows (demand for the condominiums), and finally a consequence will occur (profit.)

The consequence resulting from a combination of a decision alternative and a state of nature is a **payoff**.

A **payoff table** shows payoffs for all combinations of decision alternatives and states of nature.

Payoffs can be expressed in terms of profit, cost, time, distance or any other appropriate measure.

| | State of Nature | |
|---|---|---|
| Decision Alternative | $s_1$, Strong Demand | $s_2$, Weak Demand |
| Small complex, $d_1$ | 8 | 7 |
| Medium complex, $d_2$ | 14 | 5 |
| Large complex, $d_3$ | 20 | −9 |

Because PDC wants to select the complex size that provides the largest profit, *profit* is used as the consequence. The payoff table shows the profits expressed in millions of dollars.

The payoff associated with decision alternative *i* and state of nature *j* will be denoted as $V_{ij}$.

For example, $V_{ij} = 20$ indicates that a payoff of $20 million if the decision is to build a large complex ($d_3$) and the strong demand state of nature ($s_1$) occurs.

# 19.1 Decision Tree

A **decision tree** is a graphical and chronological representation of the decision problem. Each decision tree has two types of **nodes**:

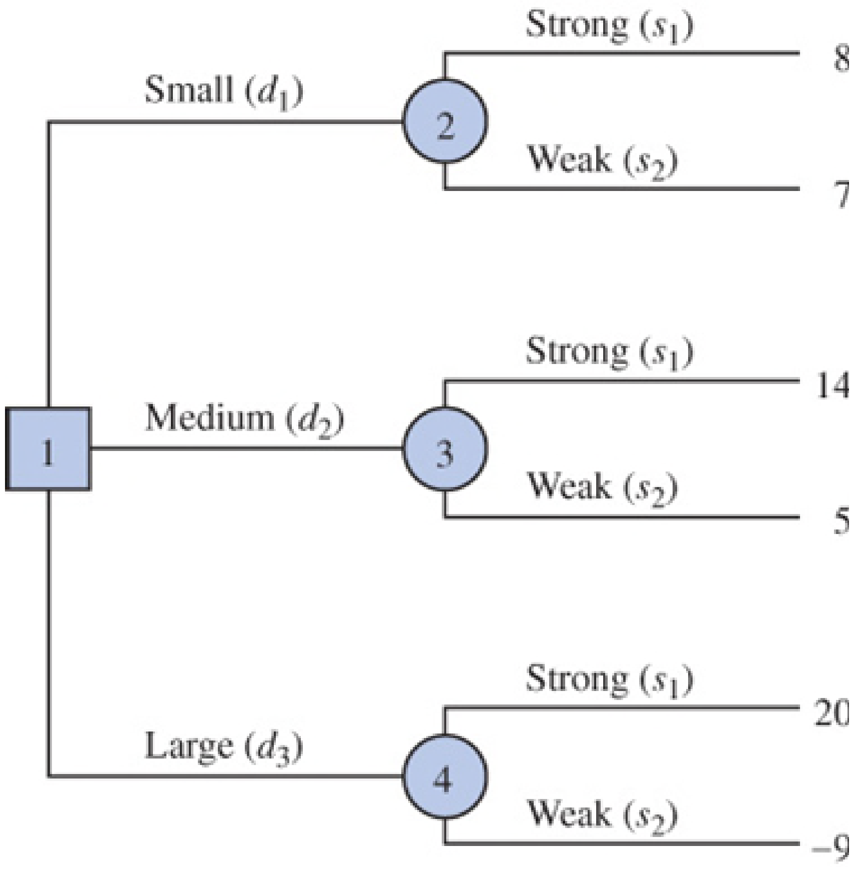
- *Square nodes*: decision nodes corresponding to the decision alternatives.
- *Round nodes*: chance nodes corresponding to the states of nature.

The decision tree for PDC has four nodes, numbered 1-4. Node 1 is a decision node, and nodes 2-4 are chance nodes.

The **branches** leaving each round node represent the different states of nature.

The branches leaving each square node represent the different decision alternatives.

At the end of each limb of the decision tree are the payoffs associated with each combination of a decision alternative and a state of nature.




CENGAGE

# 19.2 Decision Making with Probabilities

If probabilistic information regarding the states of nature is available, one may use the **<u>expected value (EV) approach</u>**.

Let us first define the EV of a decision alternative $d_i$ as:

$$\boldsymbol{EV(d_i) = \sum_{j=1}^{N} P(s_j)V_{ij}}$$

Where:

$N$ = the number of states of nature

$P(s_j)$ = the probability of state of nature $s_j$ *with* $\sum P(s_j) = 1$.

In other words, the expected value of a decision alternative is the sum of weighted payoffs.

The weight for a payoff is the probability of the associated state of nature and therefore the probability that the payoff will occur.

# 19.2 Expected Value

Let us suppose PDC has optimistically assigned a probability of 0.8 to the state of nature that demand for condominiums is strong, and a probability of 0.2 that demand is weak.

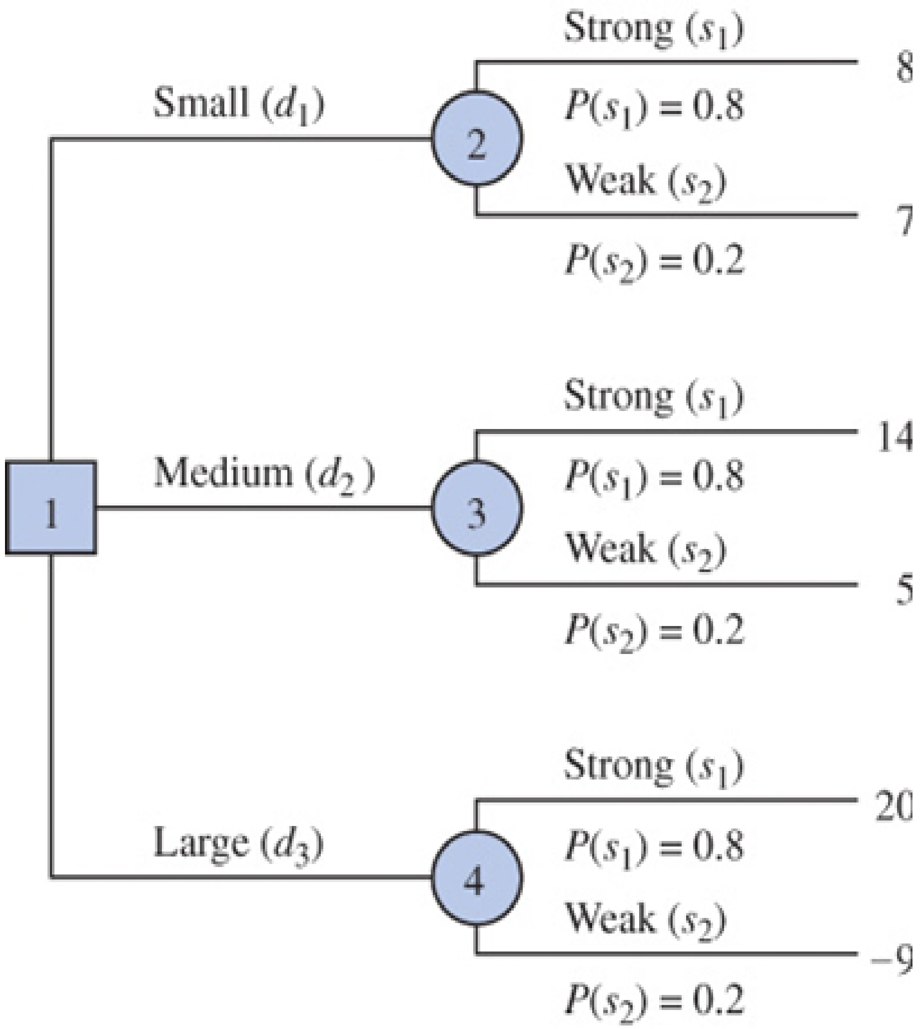
Using the payoff table and the EV equation, we compute the expected value for each of the three decision alternatives as

$$EV(d_1) = 0.8(8) + 0.2(7) \quad = 7.8$$

$$EV(d_2) = 0.8(14) + 0.2(5) \quad = 12.2$$

$$EV(d_3) = 0.8(20) + 0.2(-9) = 14.2$$

Thus, using the expected value approach, we find that the recommended decision is the large condominium complex, with $EV = \$14.2$ million.



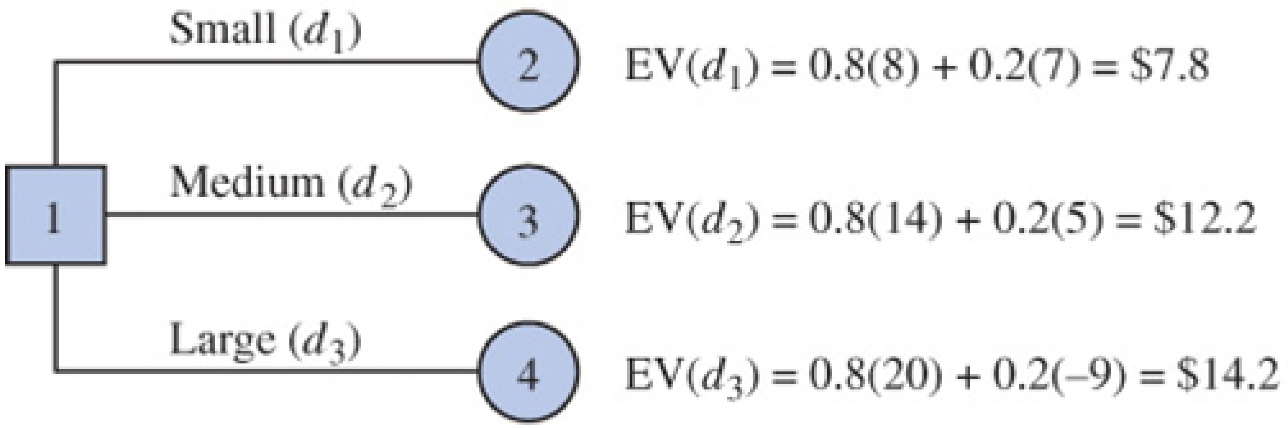

CENGAGE

# 19.2 The Expected Value Approach to the PDC Problem

The calculations required to identify the decision alternative with the best expected value for the PDC problem can be conveniently carried out on a decision tree, as shown here.

First, we compute the expected value at each chance node.

Then, working backward through the decision tree, we weight each possible payoff by its probability of occurrence, and obtain the expected values for nodes 2, 3, and 4.




Because the decision maker controls the branch leaving decision node 1, and because we are trying to maximize the expected profit, it follows that for PDC the best decision alternative at node 1 is $d_3$, with an expected value, $EV = \$14.2$ million.

CENGAGE

# 19.2 Expected Value of Perfect Information (EVPI)

Information to improve the probability estimates for the states of nature is often available.

The **expected value of perfect information (EVPI)** is the increase in the expected profit that would result if one knew with certainty which state of nature will occur.

The EVPI provides an upper bound on the expected value of any sample or survey information, and is computed as follows:

$$\text{EVPI} = |\text{EVwPI} - \text{EVwoPI}|$$

Where

EVPI = expected value of perfect information

EVwPI = expected value *with* perfect information about the states of nature

EVwoPI = expected value *without* perfect information about the states of nature

Note the role the absolute value plays in the EVPI equation: for minimization problems, EVwPI is always less than or equal to EVwoPI.

CENGAGE

# 19.2 EVPI for the PDC Problem

PDC's optimal decision strategy when the perfect information becomes available is stated as follows:

If $s_1$, select $d_3$ with a payoff of \$20 million

If $s_2$, select $d_1$ with a payoff of \$7 million

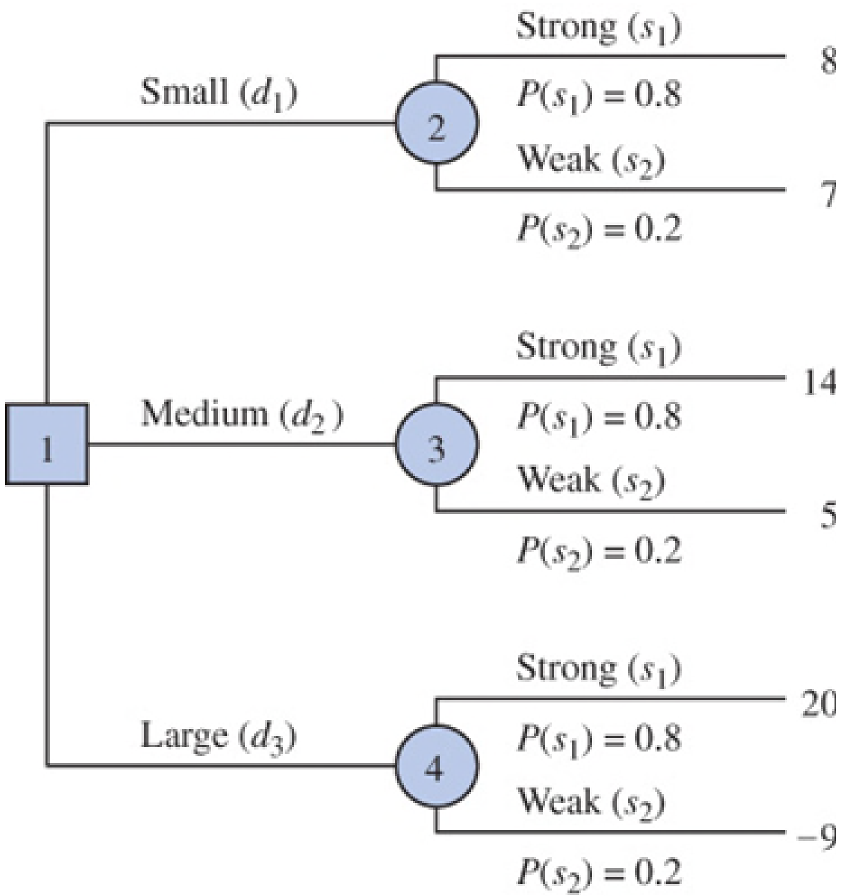
It follows that:

$$\text{EVwPI} = 0.8(20) + 0.2(7) = \$17.4 \text{ million}$$

$$\text{EVwoPI} = 0.8(20) + 0.2(-9) = \$14.2 \text{ million}$$

$$\text{EVPI} = |\text{EVwPI} - \text{EVwoPI}| = |17.4 - 14.2| = \$3.2 \text{ million}$$

Thus, \$3.2 million represents a theoretical upper bound to the expected value of sample information.




A market research study will never provide "perfect" information. However, in a good market research study, the information gathered might be worth a sizable portion of the \$3.2 million. Given the EVPI of \$3.2 million, PDC might seriously consider a market survey.

# 19.3 Decision Analysis with Sample Information

Frequently, decision makers have preliminary or **prior probability** assessments for the states of nature that are the best probability values available at that time.

However, to make the best possible decision, the decision maker may want to seek additional information about the states of nature.

This new information, often obtained through experiments designed to provide **sample information**, can be used to revise the prior probabilities so that the final decision is based on more accurate probabilities for the states of nature.

The revised probabilities are called **posterior probabilities**.

PDC believes that a 6-month market research study, designed to learn more about the condominium market acceptance, will provide a report with one of the following two outcomes:

Favorable: several individuals expressed interest in purchasing a condominium.

Unfavorable: very few individuals expressed interest in purchasing a condominium.

# 貝氏計算(Bayesian Computing)

- 什麼是貝氏計算(Bayesian Computing)

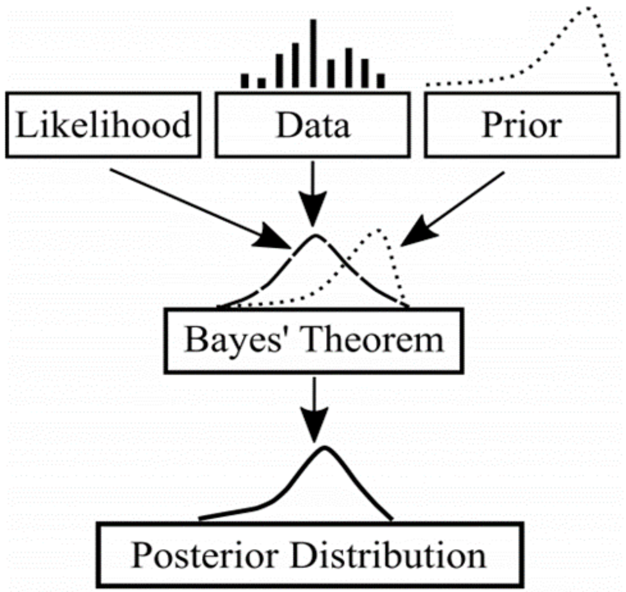
→根據新蒐集的資訊，修正更新對某個未知事物發生的機率預測。

- 貝氏方法視經驗為先驗資訊(Prior Information)，結合本次蒐集的資料(Data)，綜合可得新的經驗(驗後結果, Posterior)：

**先驗資訊 ＋ 實驗結果 → 驗後結果**

**(Prior ＋ Data → Posterior)**

https://medium.com/@ram420/bayes-theorem-with-conditional-probability-793bf9caba92




https://www.the-tls.com/regular-features/footnotes-to-plato/thomas-bayes-science-crisis

# 協助我們更理性的判斷

A patient takes a lab test and the result comes back positive. The test returns a correct positive result in 99% of the cases in which the disease is actually present, and a correct negative result in 98% of the cases in which the disease is not present. Furthermore, .001 of all people have this cancer.

https://tomrocksmaths.com/2021/08/31/bayes-theorem-and-disease-testing/


$$P(\text{cancer}) = .001 \qquad P(\sim \text{cancer}) = .999$$

$$P(+ \mid \text{cancer}) = .99 \qquad P(- \mid \text{cancer}) = .01$$

$$P(+ \mid \sim\text{cancer}) = .02 \qquad P(- \mid \sim\text{cancer}) = .98$$

$$P(\text{cancer} \mid +) = \frac{P(+ \mid cancer)\, P(cancer)}{P(+)} = \mathbf{.047}$$

# 可能的計算方式

- 假設某地區有一百萬人：

  → 999,000人健康，1,000人罹患癌症

  → 檢查出陽性反應者：

  (1) 健康者中有 $999{,}000 \times 2\% = 19{,}980$

  (2) 癌症患者中有 $1{,}000 \times 99\% = 990$

因此，陽性反應者中罹患癌症的比例：

https://dlsun.github.io/probability/bayes.html


$$P(cancer \mid +) = \frac{990}{19{,}980 + 990} = \frac{990}{20{,}970} \cong 4.72\%$$

# 19.3 Decision Tree with Sample Information

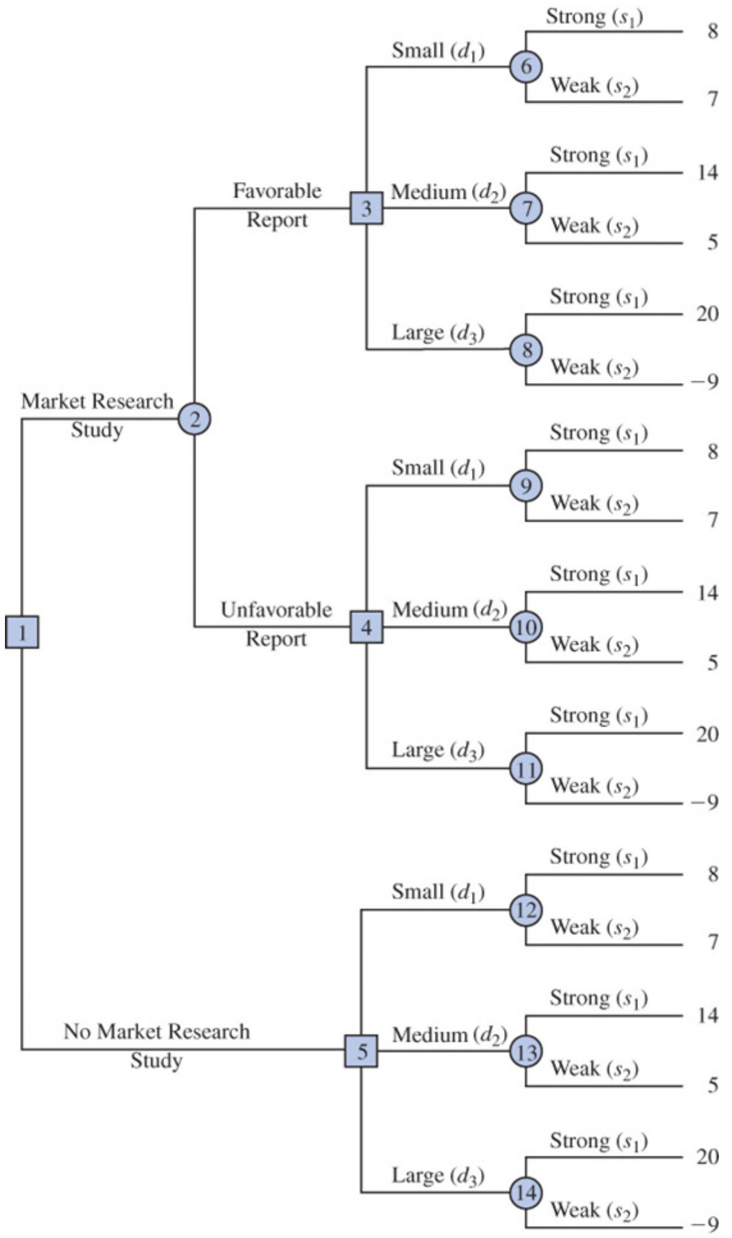
PDC must decide about whether the market research should be conducted, and the size of condominium project to build, in case:

- of favorable report.
- of unfavorable report.
- the market research study is not conducted.

Thus, the decision tree is composed by four decision nodes:

- Node 1: whether to conduct the market research study.
- Node 3: whether to build the complex if report is favorable.
- Node 4: whether to build the complex if report is unfavorable.
- Node 5: whether to build the complex if no market research study is conducted.

Chance node 2 indicates the outcomes of the market research report, and chance nodes 6-14 whether demand will be strong or weak.

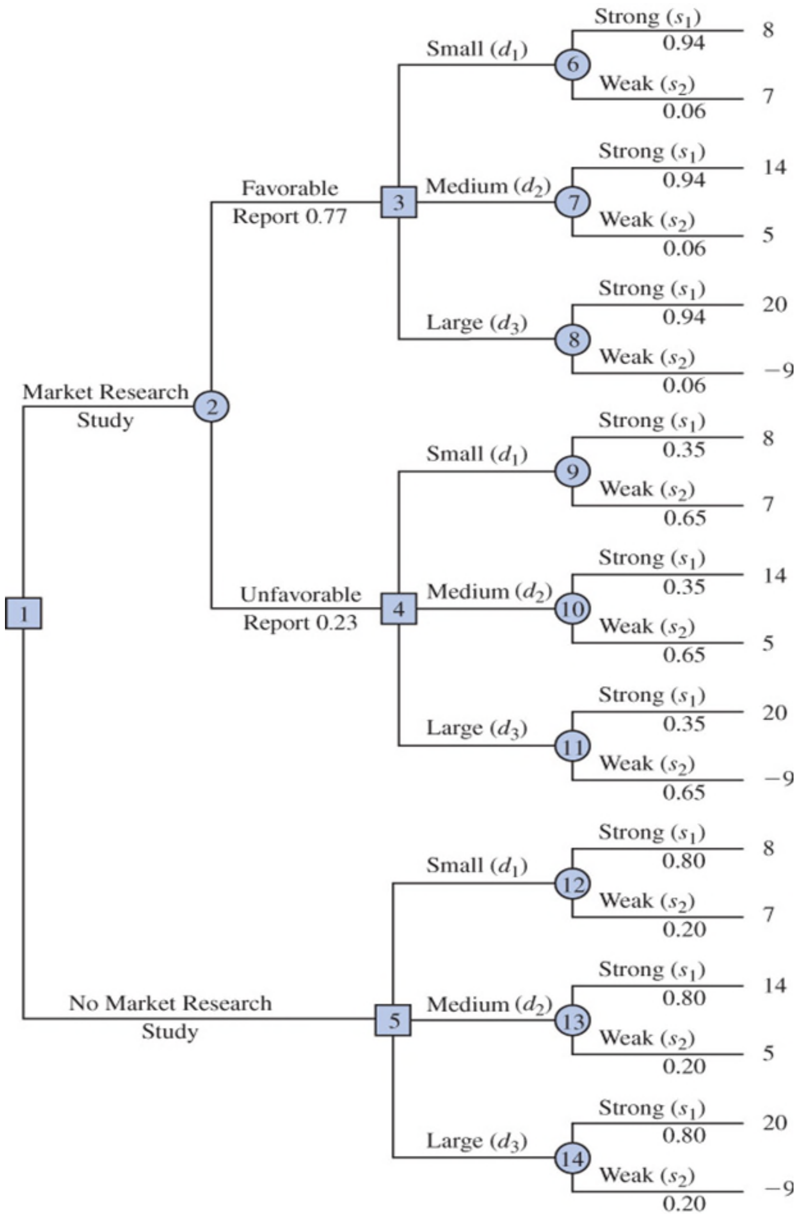



CENGAGE

# 19.3 Decision Tree Analysis

Based on history about similar past market research surveys, PDC has developed the following branch probabilities.

- Whether the market research study is undertaken
  - $P(\text{Favorable report}) = 0.77$
  - $P(\text{Unfavorable report}) = 0.23$
- Combinations of market research outcome and state of nature:
  - $P(\text{Strong demand} \mid \text{favorable report}) = 0.94$
  - $P(\text{Weak demand} \mid \text{favorable report}) = 0.06$
  - $P(\text{Strong demand} \mid \text{unfavorable report}) = 0.35$
  - $P(\text{Weak demand} \mid \text{unfavorable report}) = 0.65$
- Prior probabilities apply in case of no market research:
  - $P(\text{Strong demand}) = 0.8$
  - $P(\text{Weak demand}) = 0.2$




CENGAGE

# 19.3 Decision Strategy

A **decision strategy** is a sequence of decisions and chance outcomes where the decisions chosen depends on the yet-to-be-determined outcomes of chance events.

The approach used to determine the optimal decision strategy is based on a backward pass through the decision tree using the following steps:

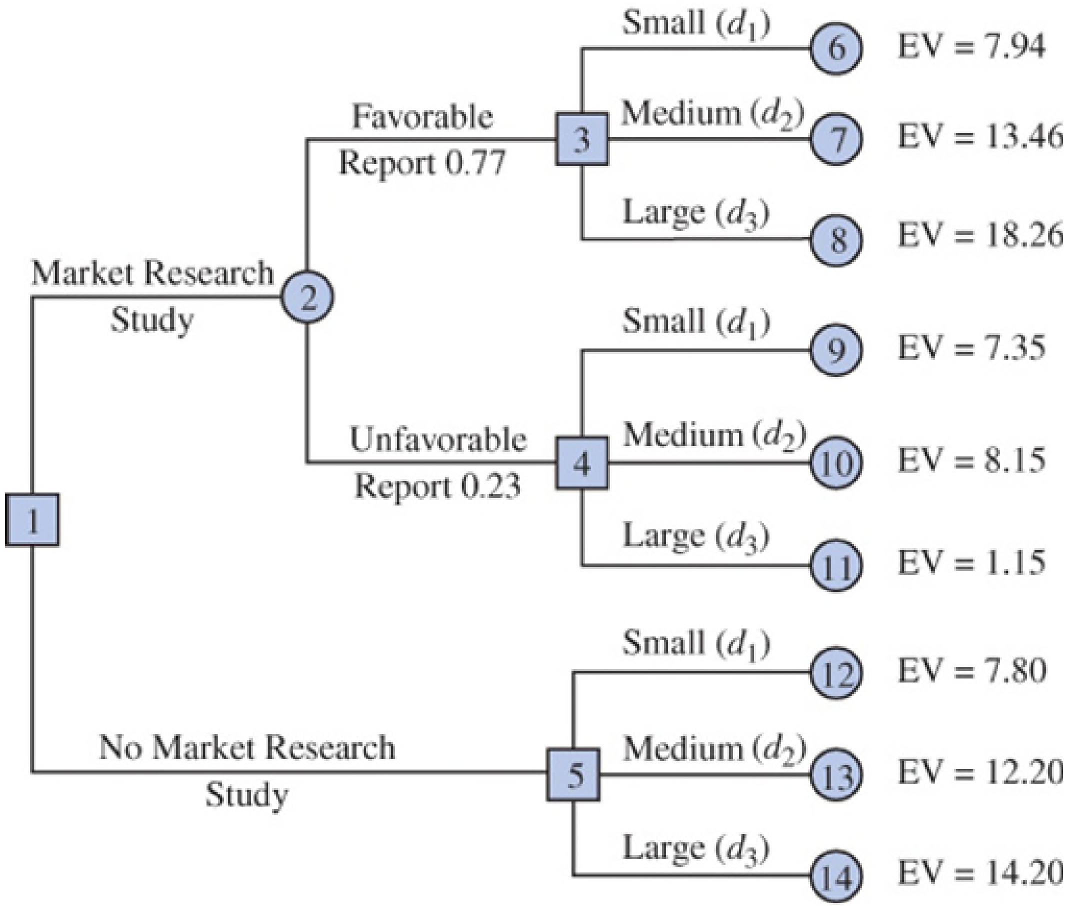
1. At chance nodes, compute the expected value by multiplying the payoff at the end of each branch by the corresponding branch probabilities.
2. At decision nodes, select the decision branch that leads to the best expected value. This expected value becomes the expected value at the decision node.

Let us determine the optimal decision strategy for the PDC condominium project problem by first considering the backward pass calculations at each of the decision tree terminal chance nodes.

# 19.3 Decision Strategy: Chance Nodes

The backward pass calculations computed at chance nodes 6 to 14 provide the following expected values:

$$\text{EV(Node 6)} = 0.94(8 + 0.06(7)) = 7.94$$
$$\text{EV(Node 7)} = 0.94(14 + 0.06(5)) = 13.46$$
$$\text{EV(Node 8)} = 0.94(20 + 0.06(-9)) = 18.26$$
$$\text{EV(Node 9)} = 0.35(8 + 0.65(7)) = 7.35$$
$$\text{EV(Node 10)} = 0.35(14 + 0.65(5)) = 8.15$$
$$\text{EV(Node 11)} = 0.35(20 + 0.65(-9)) = 1.15$$
$$\text{EV(Node 12)} = 0.80(8 + 0.20(7)) = 7.80$$
$$\text{EV(Node 13)} = 0.80(14 + 0.20(5)) = 12.20$$
$$\text{EV(Node 14)} = 0.80(20 + 0.20(-9)) = 14.20$$




CENGAGE

# 19.3 Decision Strategy: Decision Nodes

For each of the decision nodes, we select the decision alternative branch that provides the best (largest) EV.

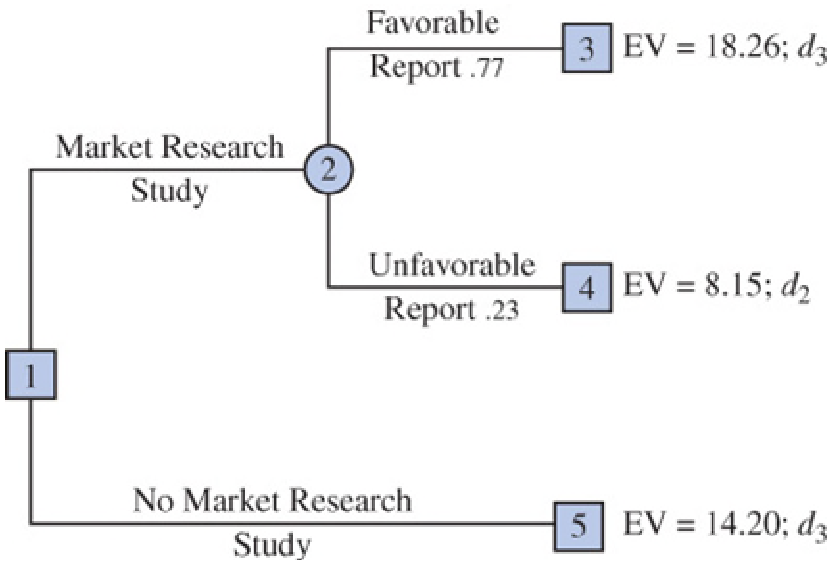
- At node 3 (favorable report), $\text{EV}(d_3) = 18.26$, is greater than $\text{EV}(d_1) = 7.94$ and $\text{EV}(d_2) = 13.46$.
- A similar decision at node 4 (unfavorable report) yields $\text{EV}(d_2) = 8.15$.
- At node 5 (no market research study) the best alternative branch has $\text{EV}(d_3) = 14.20$.

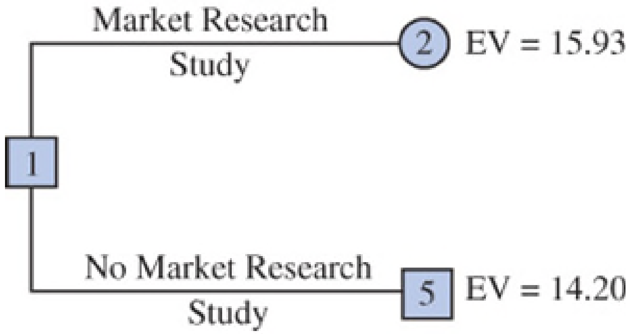
The EV at chance node 2 can now be computed as

$$\text{EV(Node 2)} = 0.77\ \text{EV(Node 3)} + 0.23\ \text{EV(Node 4)}$$

$$\text{EV(Node 2)} = 0.77(18.26) + 0.23(8.15) = 15.93$$

Finally, at node 1, the EV for node 2 is larger than the EV for node 5. Thus, we have, $\text{EV(Node 1)} = 15.93$.



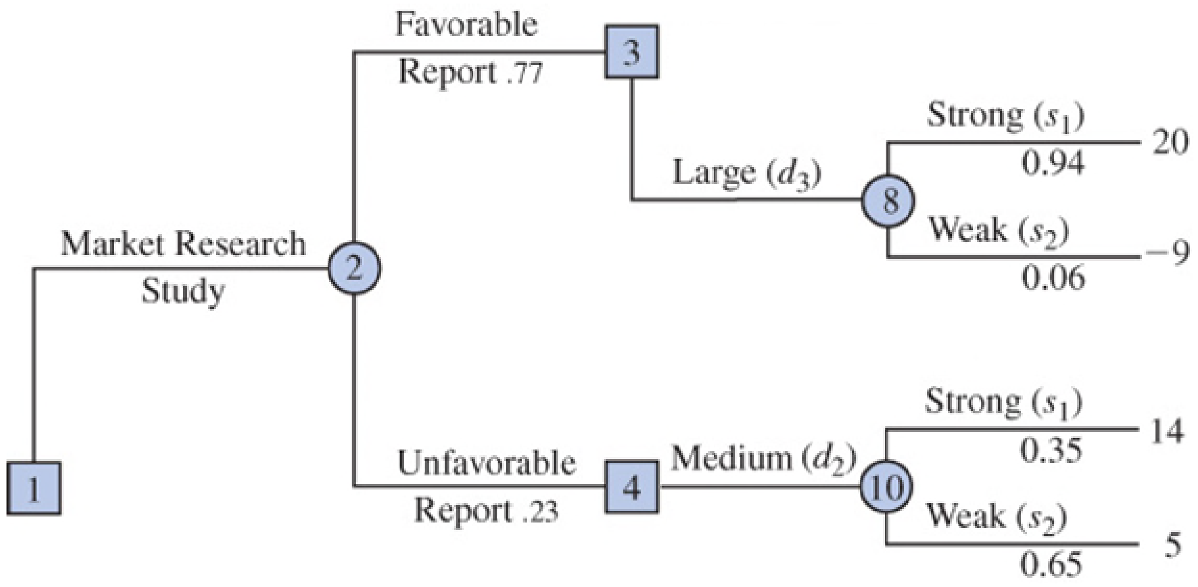




CENGAGE

# 19.3 Optimal Decision Strategy for the PDC Problem

The optimal decision for PDC is to conduct the market research study and then carry out the following decision strategy:

- If the market research report is favorable, construct the large condominium complex.
- If the market research report is unfavorable, construct the medium condominium complex.




The steps exemplified in the analysis of the PDC problem with sample information can be used to analyze more complex sequential decision problems.

# 19.3 Expected Value of Sample Information (EVSI)

The **expected value of sample information (EVSI)** is the additional expected profit possible through knowledge of the sample or survey information, and is computed as follows

$$\text{EVSI} = |\text{EVwSI} - \text{EVwoSI}|$$

Where,

EVwSI is the expected value with sample information

EVwoSI the expected value without sample information

Because EVSI cannot exceed EVPI, we can define the **efficiency of sample information** as

$$E = \frac{EVSI}{EVPI} \times 100$$

CENGAGE

# 19.3 EVSI for the PDC Problem

In the PDC problem, we have

$EVwSI = 15.93$ (associated with the market research study)

$EVwoSI = 14.20$ (in case the market research study is not undertaken)

Thus, we can conclude that

$$\text{EVSI} = |\text{EVwSI} - \text{EVwoSI}| = |15.93 - 14.20| = \$1.73 \text{ million.}$$

That is, conducting the market research study adds $1.73 million to the PDC expected value.

We can also calculate the efficiency of sample information as

$$E = \frac{EVSI}{EVPI} \times 100 = \frac{1.73}{3.20} \times 100 = 54.1\%$$

Thus, the information from the market research study is 54.1% as efficient as perfect information.

CENGAGE

# 19.4 Computing Branch Probabilities with Bayes' Theorem

**Bayes' theorem** can be used to compute the *branch probabilities* specified for the PDC decision tree chance nodes.

The PDC decision tree is shown again to the right, where:

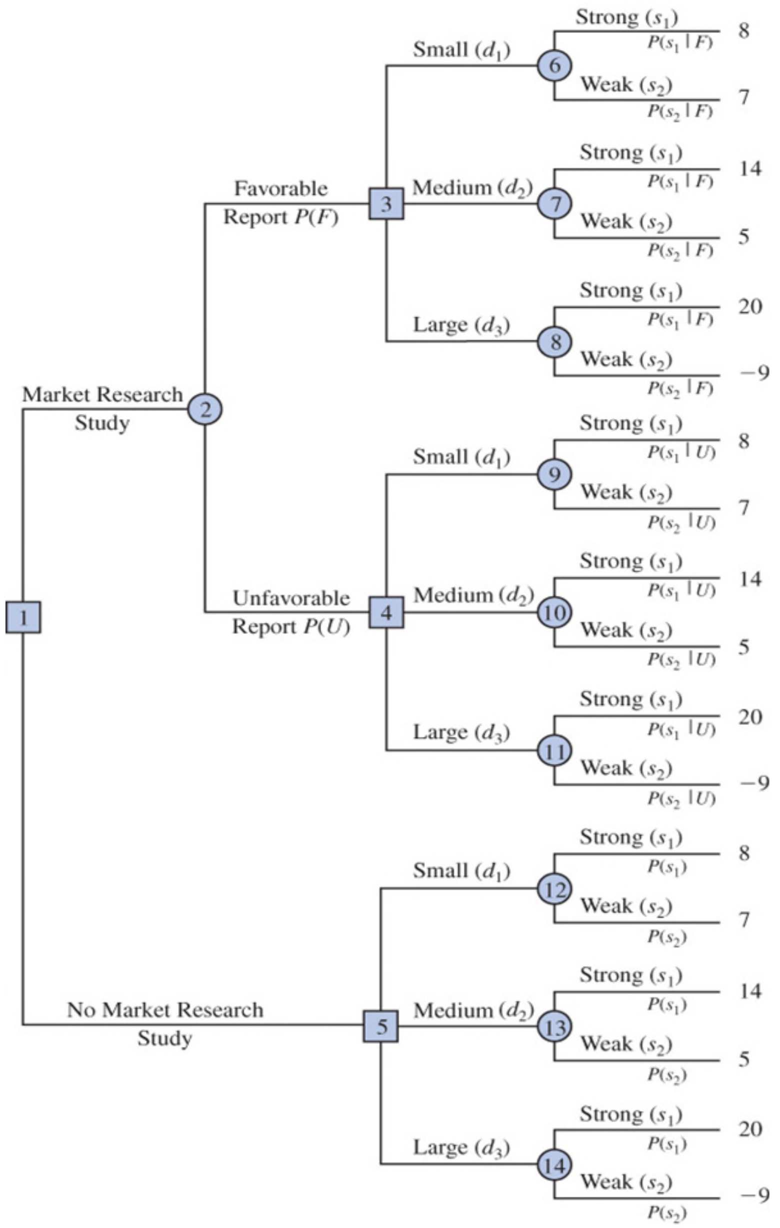
Market research indicator: *F* = favorable *U* = unfavorable

Market demand (state of nature): $s_1$ = strong $s_2$ = weak

We need to know the branch probabilities at the following nodes:

| | | | |
|---|---|---|---|
| Node 2: | $P(F)$ | and | $P(U)$ |
| Nodes 6-8: | $P(s_1 \vert F)$ | and | $P(s_2 \vert F)$ |
| Nodes 9-11: | $P(s_1 \vert U)$ | and | $P(s_2 \vert U)$ |
| Nodes 12-14: | $P(s_1)$ | and | $P(s_2)$ |

$P(s_1|F)$, $P(s_2|F)$, $P(s_1|U)$, and $P(s_2|U)$ are referred to as **posterior probabilities** because they are conditional probabilities based on the outcome of the sample information.




CENGAGE

# 19.4 Estimating Conditional Probabilities

To calculate the posterior probabilities, we must know the **conditional probability** of the market research outcomes (the sample information) *given* each state of nature.

The conditional probabilities may be estimated via the historical frequency of market research reports in cases where a given demand was ultimately observed.

In the PDC problem we assume that the following assessments are available for these conditional probabilities:

| | Market Research | |
|---|---|---|
| State of Nature | Favorable, $F$ | Unfavorable, $U$ |
| Strong demand, $s_1$ | $P(F \mid s_1) = 0.90$ | $P(U \mid s_1) = 0.10$ |
| Weak demand, $s_2$ | $P(F \mid s_2) = 0.25$ | $P(U \mid s_2) = 0.75$ |

Note how the probability of a favorable market report is much higher if demand turned out to be strong, and the opposite is true in case of an unfavorable market report.

The conditional probability $P(F|s_2)$ is often referred to as a false positive, and $P(U|s_1)$ as a false negative.

CENGAGE

# 19.4 Computing Posterior Probabilities

To compute the posterior probabilities for the PDC problem on a favorable market research report ($F$) that are shown in the top table of the next slide, follow these steps:

**Step 1**. In column 1, enter the states of nature. In column 2, enter the *prior probabilities* for the states of nature. In column 3, enter the *conditional probabilities* of a favorable market research report ($F$) given each state of nature.

**Step 2**. In column 4, compute the **joint probabilities** by multiplying the prior probability values in column 2 by the corresponding conditional probability values in column 3.

**Step 3**. Sum the joint probabilities in column 4 to obtain the probability of a favorable market research report, $P(F)$.

**Step 4**. Divide each joint probability in column 4 by $P(F) = 0.77$ to obtain the revised or *posterior probabilities*, $P(s_1|F)$ and $P(s_2|F)$.

Repeating these steps for the unfavorable market research report ($U$) produces the posterior probabilities $P(s_1|U)$ and $P(s_2|U)$ shown in the bottom table of the next slide.

# 19.4 Posterior Probabilities for the PDC Problem

| Market Research Indicator | State of Nature $s_j$ | Prior Probability $P(s_j)$ | Conditional Probability $P(F \mid s_j)$ | Joint Probability $P(F \cap s_j)$ | Posterior Probability $P(s_j \mid F)$ |
|---|---|---|---|---|---|
| $F$ | $s_1$ | 0.80 | 0.90 | 0.72 | 0.94 |
| $F$ | $s_2$ | 0.20 | 0.25 | 0.05 | 0.06 |
| | | 1.00 | | $P(F) = 0.77$ | 1.00 |

| Market Research Indicator | State of Nature $s_j$ | Prior Probability $P(s_j)$ | Conditional Probability $P(U \mid s_j)$ | Joint Probability $P(U \cap s_j)$ | Posterior Probability $P(s_j \mid U)$ |
|---|---|---|---|---|---|
| $U$ | $s_1$ | 0.80 | 0.10 | 0.08 | 0.35 |
| $U$ | $s_2$ | 0.20 | 0.75 | 0.15 | 0.65 |
| | | 1.00 | | $P(U) = 0.23$ | 1.00 |

CENGAGE

## Guessing the Pet

Is it a cat or a dog in the box?

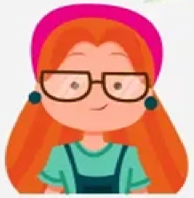
P(Dog) = 0.5

P(Cat) = 0.5

Initial Belief: 50/50 chances

## We have a CLUE

The Pet is very Quiet



→ (P(Quiet|Cat))= 80% * or * 0.8

→ (P(Quiet|Dog))= 30% * or * 0.3

# Applying Bayes Theorem to Find Probability

$$P(Cat \mid Quiet) = \frac{P(Quiet\ Cat) \times P(Cat)}{P(Quiet)} = 72.7\%$$

$$P(Dog \mid Quiet) = \frac{P(QuietDog) \times P(Dog)}{P(Quiet)} = 27.3\%$$

https://www.geeksforgeeks.org/maths/bayes-theorem/

# Summary

- Decision analysis can be used to determine an optimal decision strategy when a decision maker is faced with decision alternatives and an uncertain pattern of future events.
- The goal of decision analysis is to identify the optimal decision strategy given information about the uncertain events and the possible consequences or payoffs.
- The uncertain future events are called chance events and the outcomes of the chance events are called states of nature
- We showed how payoff tables and decision trees can be used to structure a decision problem and describe the relationships among the decisions, the chance events, and the consequences.
- We also introduced the concepts of perfect information and sample information.
- We demonstrated how Bayes' theorem can be used to compute the branch probabilities for decision trees and to update prior probabilities with revised posterior probabilities.

CENGAGE