probability of the other; if they are dependent, the probability of the occurrence of one event is affected by the occurrence or nonoccurrence of the other. The following equations, derived from the laws of probability, are relevant to conditional probability and the derivation of Bayes' theorem (Freund and Williams 1977).

The formula for independent events is derived from the *special multiplication rule*

$$P(A \cap B) = P(A) \cdot P(B)$$

The formula for dependent events is derived from the *general multiplication rule*

$$P(A \cap B) = P(B) \cdot P(A|B)$$

Because it does not matter which event is referred to as A or B, we can say that

$$P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$$

which is Bayes' theorem that deals specifically with dependent events and allows for the updating of probability values should new information arise. In the previous equation for the *general multiplication rule*, we see the conditional probability $P(A|B)$. This is defined as the probability that event A will occur given that event B has occurred. This conditional probability is where the new information would be applied to revise the posterior probability.

8.4.3 Bayes' Theorem Derived from Conditional Probability

To illustrate how Bayes' theorem can be derived simply from conditional probability, let us consider the following example:

> In a given year, 80 civil engineering students at a local university apply for a summer internship at a prestigious engineering firm. Of all the applicants, 35 have prior internship experience, whereas the rest do not. Of the 50 applicants considered to be upper-class students (juniors and seniors), 20 have prior internship experience. There is only one position available. What is the probability that an upper-class student, given that they had no prior experience, will be hired for the open position?

The data given can be tabulated as shown in Table 8-1, where U denotes the selection of an upper classman, L denotes the selection of a lower classman,

Table 8-1. Initial Table for Conditional Probability Example

	Experienced (E)	No experience (N)	Total
Upper Classmen (U)	20	—	50
Lower Classmen (L)	—	—	—
Total	35	—	80

E denotes the selection of a student with prior experience, and N denotes the selection of a student with no experience. Using the given data, we can fill the remaining cells with a bit of arithmetic, as shown in Table 8-2.

The individual probability of selecting a specific student applicant can now be determined by making the following calculations. The probability of selecting an upper classman from the entire sample is

$$P(U) = \frac{20 + 30}{80} = 0.625$$

The probability of selecting a lower classman is

$$P(L) = \frac{15 + 15}{80} = 0.375$$

The probability of selecting a student with prior internship experience is

$$P(E) = \frac{20 + 15}{80} = 0.4375$$

The probability of selecting a student with no prior internship experience is

$$P(N) = \frac{30 + 15}{80} = 0.5625$$

The probability of selecting an upper classman and a student with no experience is

$$P(U \cap N) = \frac{30}{80} = 0.375$$

Therefore, to determine the probability that an upper classman is selected, if given that he or she has no prior experience and assuming that each applicant has an equal chance $(1/80)$ of being hired, we compute

$$P(U|N) = \frac{30}{45} = 0.667 = 66.7\%$$

Table 8-2. Final Table for Conditional Probability Example

	Experienced (E)	No experience (N)	Total
Upper Classmen (U)	20	30	50
Lower Classmen (L)	15	15	30
Total	35	45	80

Note that this conditional probability can be also written as

$$P(U|N) = \frac{30/80}{45/80} = \frac{P(U \cap N)}{P(N)} = \frac{30}{45} = 0.667$$

Therefore,

$$P(U \cap N) = P(U|N) \cdot P(N)$$

Similarly, $P(N|U) \cdot P(U) = P(N \cap U)$. Because $P(U \cap N) = P(N \cap U)$, we can equate the two formulas. Thus,

$$P(U|N \cdot (P(N)) = P(N|U) \cdot P(U)$$

And, in general,

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

which is Bayes' theorem.

8.5 Logic of Prior/Posterior Probabilities and Tree Diagrams

The logic of Bayes' theorem can be applied in either direction. The first logic to consider would be from *cause to effect*, as shown in Fig. 8-5. Event A represents the cause, and event B represents the effect, or outcome. The probabilities that follow this logic are the a priori, or prior, probability and its respective conditional probabilities. The a priori probabilities are used to derive the a posteriori probabilities. Conversely, reasoning in the opposite direction would go from *effect to cause*, using the a posteriori, or posterior, probability and its conditional probabilities as a check to confirm the a priori probability values. This *effect to cause* logic, applied to the example in Fig. 8-5, can now be seen in Fig. 8-6, where events B and B' are now the probable causes, and events $A, A_1 \ldots A_k$ are now the probable effects. Generally, when going from effect to cause, the a posteriori becomes the new a priori. Whether one starts with a priori or a posteriori is often a function of the type of data available, but they are both mutually interrelated.

Tree diagrams (Figs. 8-5 and 8-6) are used in conjunction with Bayes' theorem for multiple, mutually exclusive events. When events are *mutually exclusive*, it means that the events cannot occur at the same time. As shown in Fig. 8-5, events $A, A_1 \ldots A_k$ are the mutually exclusive causes, where A is one event, A_1 is the second event, and A_k represents a subsequent event, and events B and B' are the mutually exclusive effects/outcomes. Conversely, in Fig. 8-6, events B and B' are the mutually exclusive causes, and events $A, A_1 \ldots A_k$ are the mutually exclusive effects. The prime shown in event B' indicates its distinction and mutual exclusivity from B.

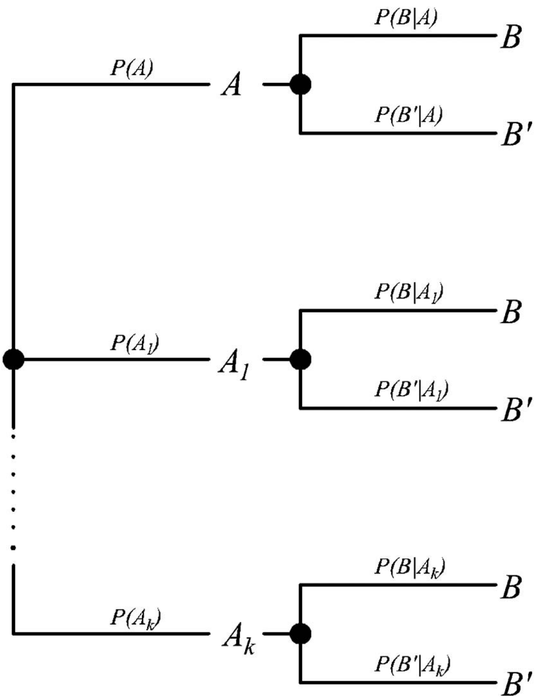


Fig. 8-5. General/prior (a priori) tree diagram

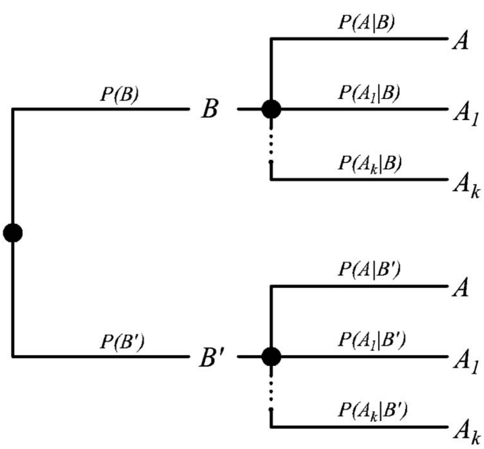


Fig. 8-6. Posterior (a posteriori) tree diagram

These trees are a useful aid when visualizing the logic behind the processes, especially when considering more than two causes and effects. Hence, this chapter builds on earlier chapters that used trees (Ch. 6) and cause-effect diagrams (Ch. 4). All together, these three chapters complement each other.

There are three characteristics to consider about these tree diagrams:

1) Branches stem from a central node (shown as a dot on the tree diagrams).
2) The total probability of all of the branches emanating from a node must be equal to 1.0, or 100%.
3) The events assigned to branches stemming from a node must be mutually exclusive.

The general formula for Bayes' theorem, when considering multiple, mutually exclusive events, is an expansion of what was previously derived:

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \cdots + P(A_k) \cdot P(B|A_k)}$$

for $i = 1, 2, \cdots, k$.

where $P(A_1) \cdot P(B|A_1)$ is equal to the joint probability of reaching B from A_l; $P(A_2) \cdot P(B|A_2)$ is equal to the joint probability of reaching B from A_2; and $P(A_k) \cdot P(B|A_k)$ is equal to the joint probability of reaching B from A_k. The entire denominator is equal to the sum of all the joint probabilities, which is $P(B)$ in the above formula.

Tabulating these joint probabilities further aid in their calculation. The use and application of these tree diagrams and prior and posterior probability calculation tables are shown in Sections 8.6.1 and 8.6.2.

A caveat to consider when calculating the prior and posterior probabilities in the tables is the decimal rounding of these probabilities. Probabilities may or may not result in the exact calculated values, due to errors in rounding. An example of this situation also is shown in Sections 8.6.1 and 8.6.2. Keep in mind that the total probability of events in a given sample space must equal 1, or 100%.

8.6 Examples and Sample Exercises: Applications of Bayes' Theorem

Now that Bayes' theorem has been derived, the following sections cover the application of Bayes' theorem through a series of simple examples of real-world situations. These sections also cover the setup and use of the a priori and a posteriori probability calculation tables.

In the first example, a client firm seeks probabilistic information on a particular contractor for an upcoming project. The contractor has a great reputation and does quality work, but has projects that run over budget and are delayed. The client firm wants detailed information on this contractor to be able to compare with other contractors. In this scenario, Bayes' theorem can help in the decision-making process by enabling the quantification and comparison of performance attributes, thereby helping to reduce risk.

The second example deals with production, management, and industrial applications. Again, Bayes' theorem can prove useful in risk assessment and decision making. In this scenario, a firearms manufacturer outsources several ancillary parts for its rifle. Although firearms are built within tight tolerances, no system is truly perfect; therefore, the parts are susceptible to manufacturing defects. Knowing the chances of defective parts occurring and being able to quantify the probability of a particular part having an increased risk of defects would be beneficial to the manufacturer. In other words, having an idea of what parts could be defective can help the manufacturer decide which parts need to have an increase in inventory, thereby reducing the risk of costs caused by delayed production because of having an insufficient quantity of a particular part on hand for a complete rifle.

8.6.1 Example 1: Contractor Performance Information

A client firm needs to decide on whether they should consider to hire a particular contractor for an upcoming project. For this, they need probabilistic information on pssible delays and budget performance. The contractor, All-Win Construction, has a good reputation, produces quality work, and typically complete their projects, within budget, at a rate of 95%. Past information reveals that the probability at which All-Win Construction encounters a delay, given that they complete a project within budget is 10%. On the other hand the probability that All-Win Construction encountered a delay, given that they went over budget is 70%.

Now, if All-Win Construction were to encounter a delay, what is the probability that they will complete a project within budget?[2] Help the client firm arrive at this information.

The first step, therefore, is to assign variables to the given events:

- C = Completed projects within budget
- N = Completed projects over budget (not within budget)
- D = Encountered delays
- ND = Encountered no delays

From here, a tree diagram can be created to organize the events given in the problem (Fig. 8-7). Each branch represents an event, has its own probability, and is typically read from left to right.

Take branch C in Fig. 8-7 as an example. The probability, or a priori, of branch C is $P(C)$. Branches D and ND stem from branch C. Subsequently, going from branch C to branch D yields a conditional probability, $P(D|C)$; correspondingly, going from branch C to branch ND yields $P(ND|C)$.

[2]This can be rephrased: What is the probability that the contractor will complete within budget, given that a delay was encountered?

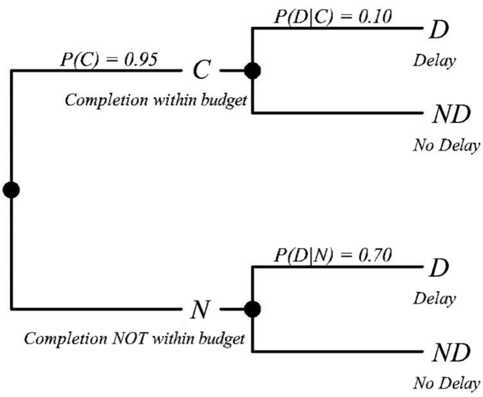


Fig. 8-7. Initial tree diagram with given events for Example 1

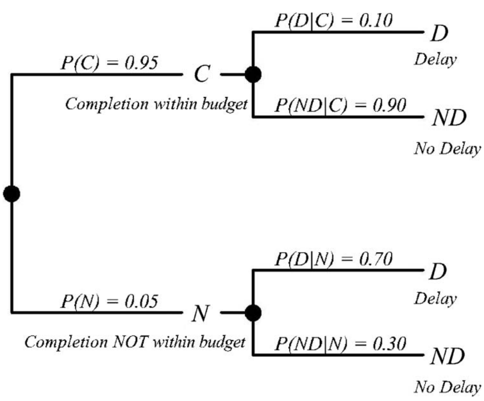


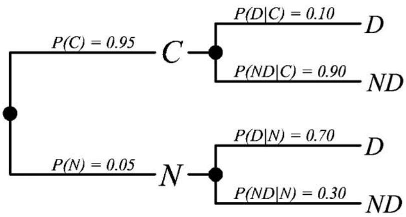
Fig. 8-8. Final (a priori) tree diagram for Example 1

Fig. 8-8 shows the tree diagram with the given probabilities, along with the complementary probability of each corresponding branch. The total probabilities of corresponding branches must equal 1, or 100%. For example, in Fig. 8-8, branch C splits off into branches D and ND. The conditional probability for branch ND is $P(ND|C) = 0.90$, because the conditional probability given for branch D is $P(D|C) = 0.1$. It is thus seen that $P(D|C) + P(ND|C) = 1 \Rightarrow P(ND|C) = 1 - P(D|C) = 0.90$. The same applies to the branches stemming from branch N.

The individual probabilities of branches D and ND can now be calculated.

Fig. 8-9 shows the calculation of these probabilities; the last row of the table is simply the sum of the probabilities for each branch. Keep in mind that the sum of these two probabilities must equal 1. This figure extends Fig. 8-8 to simulate the probabilities of D and ND. Coming from $P(D|C)$, there can be no ND possible; coming from $P(ND|C)$, there can be no D possible, and so forth. This is reflected in Fig. 8-9.

Now, Bayes' theorem can be applied to determine the a posteriori probability that the contractor will complete the project within budget, given that they encountered a delay, or $P(C|D)$.



	P(D)	P(ND)
	$P(D \mid C) \times P(C) = (0.10)(0.95) = 0.095$	---
	---	$P(ND \mid C) \times P(C) = (0.90)(0.95) = 0.885$
	$P(D \mid N) \times P(N) = (0.70)(0.05) = 0.035$	---
	---	$P(ND \mid N) \times P(N) = (0.30)(0.05) = 0.015$
$\Sigma =$	0.13	0.87

Fig. 8-9. Example 1 probability calculations for a priori $P(D)$ and $P(ND)$

$$P(C|D) = \frac{P(D|C) \cdot (C)}{P(D)} = \frac{0.095}{0.13} = 0.73 = \mathbf{73\%}$$

Thus, the probability that All-Win Construction will complete the project within budget, given that a delay is encountered is 73%. At this point, the client firm might take into consideration other attributes of All-Win Construction or consider another contractor because 73% is not a stellar rating.

To check these calculations, let us calculate the posterior probabilities (a posteriori) by building the tree diagram in reverse (or inside-out), as shown in Fig. 8-10. The tree is set up similarly to how it was set up previously, except for using the newly calculated probabilities for D and ND. The conditional probabilities, which were initially the a priori probabilities, can now be back-calculated using Bayes' Theorem using the data from the a priori calculations. If the calculations are all correct, the posterior probabilities of C and D should equal the prior probabilities of C and D with which we originally started. The calculations for these conditional posterior probabilities are as follows:

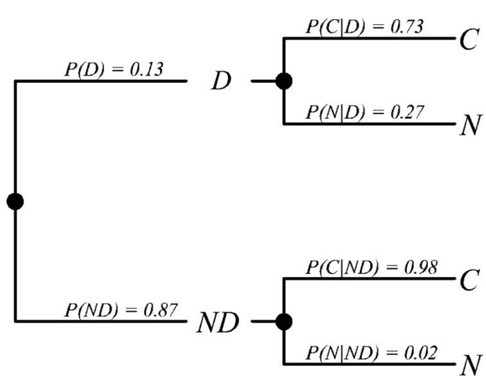


Fig. 8-10. A posteriori tree diagram for Example 1

$$P(C|D) = \frac{P(D|C) \cdot P(C)}{P(D)} = \frac{0.095}{0.13} = 0.73$$

$$P(N|D) = \frac{P(D|N) \cdot P(N)}{P(D)} = \frac{0.035}{0.13} = 0.27$$

$$P(C|ND) = \frac{P(ND|C) \cdot P(C)}{P(ND)} = \frac{0.855}{0.87} = 0.98$$

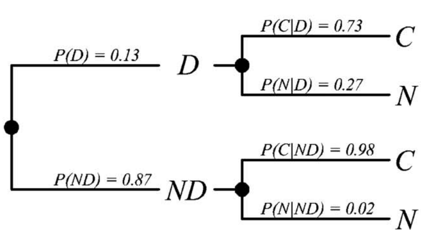
$$P(N|ND) = \frac{P(ND|N) \cdot P(N)}{P(ND)} = \frac{0.015}{0.87} = 0.02$$

Labeling the tree diagram with the appropriate values and probabilities yields the tree shown in Fig. 8-10. Once again, we see that $P(C|D) + P(N|D) = 1.0$, as is $P(C|ND) + P(N|ND)$.

As in the a priori calculations, the individual probabilities of branches C and N can now be calculated. The branches are extended in Fig. 8-11 with an aim to discover these probabilities of C and N. Observe, again, that $P(C) + P(N) = 0.95 + 0.05 = 1$. This means that the calculations are accurate. If the values are off by a considerable margin, then it is advisable to recheck the calculations; ignore or adjust minor rounding errors.

8.6.2 Example 2: Defective Parts at a Firearms Manufacturer

A firearms manufacturer orders different rifle parts from various suppliers, because it does not produce these specific parts in-house. These parts are built within strict tolerances, and defects are rare. Unfortunately, some defects do slip through quality control and end up being shipped out in the final product.



	P(C)	P(N)
	$P(C\mid D) \times P(D) = (0.73)(0.13) = 0.095$	---
	---	$P(N\mid D) \times P(D) = (0.27)(0.13) = 0.035$
	$P(C\mid ND) \times P(ND) = (0.98)(0.87) = 0.853$	---
	---	$P(N\mid ND) \times P(ND) = (0.02)(0.87) = 0.017$
$\Sigma =$	0.95	0.05

Fig. 8-11. A posteriori probability calculations for Example 1

Supplier A makes the trigger assemblies of which, records show, about 1% have been defective. Supplier B makes the rifle butt stocks, of which 3% have been defective; Supplier C makes the slings, of which 2% have been defective; and Supplier D makes the forward grips, of which 5% have been defective. The rifle manufacturer's current inventory is shown in Table 8-3.

Choosing one part at random, what is the probability that it is defective? Also, which part would have the greatest likelihood of being defective?

First, the a priori must be determined before the tree diagram can be created. To do this, start by taking the total number of units for each part and dividing that by the total number of parts (2,550 units). For example, the manufacturer has 400 units of Supplier A's trigger assemblies. Therefore, the probability of a trigger assembly being chosen at random is $400/2{,}550 = 15.7\%$.

The remaining probabilities are calculated similarly and shown in Table 8-4. The conditional probabilities can be deduced from the problem statement. For example, let $P(F|A)$ represent the probability that the part is defective, given that it is from Supplier A.

Again, before a tree diagram is created, letters must first be assigned to represent the corresponding events:

- A = Supplier A; trigger assemblies
- B = Supplier B; butt stocks
- C = Supplier C; slings
- D = Supplier D; forward grips
- F = Defective
- ND = Not Defective

Using the data that was given and the data that was calculated, the tree diagram can now be created (Fig. 8-12). Applying the probability calculations used in the

Table 8-3. Current Inventory Held by Manufacturer

Trigger assemblies (A)	Rifle butt stocks (B)	Slings (C)	Forward grips (D)
400	500	1,000	650
Total number of parts = 400 + 500 + 1,000 + 650 = 2,550			

Table 8-4. a Priori for all Suppliers

Supplier A	Supplier B	Supplier C	Supplier D
15.70%	19.60%	39.20%	25.50%