8. BAYES' RULE

Revising probabilities based upon new information.

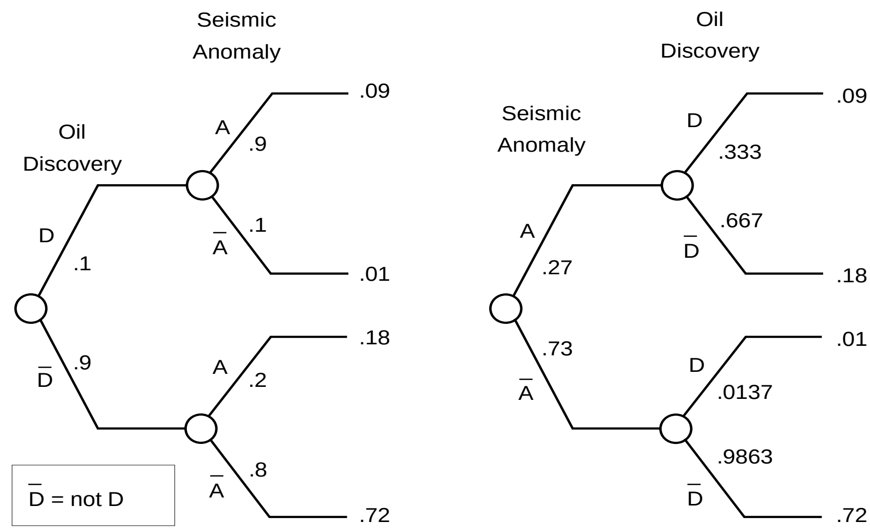
Often, our beliefs are altered by new information. In ***value of information*** problems, the analyst must understand—in advance—how additional information will affect his or her prior judgments about probabilities. Bayes' rule is a formula that enables us to recalculate probabilities based upon learning the outcomes of related events (e.g., symptoms that can be observed).

The tree form on the right is the format for asking for professional judgments about the events. The event of interest (whether there is oil) is separated from judgments about the quality of the information (second nodes). This keeps the events distinct.

Bayes' rule applies to situations characteristic of diagnostic problems. Initially, we have sufficient knowledge to establish probabilities that certain symptoms will be shown when certain conditions are present. That is, we can assess the probabilities in nature and the probability that our measurement system is reliable. Then, using Bayes' rule, we can determine the probability that the condition of interest is present when a symptom shows. In our example, seismic anomalies are symptoms of oil reservoirs.

Consider a problem in petroleum exploration. We are looking for subsurface structures that may contain oil or natural gas. A seismic survey provides a rough picture of the geology below. Obtaining seismic data provides additional, but imperfect, information. A geophysicist can estimate the probability that a seismic anomaly will (or will not) be seen given that a trap is present (or not present). The event tree on the left side of the figure shows the framework for assessing the probability judgments. The numbers these branches are the probabilities of the respective outcomes. The path probabilities at the endpoints, called joint probabilities, are the products of probabilities along each path.

The actual decision model will have the nodes *reversed*. We will learn the seismic survey outcome before we have to make the drill decision. This *inverted* structure is shown as the tree structure on the right side of the figure. In the decision model, we will have decision nodes (e.g., Drill or Drop) located between the chance nodes.



Joint Probability Table

The joint probability table shows all possible states of nature for the problem. The table presents the numbers in a convenient form. The numbers inside the table are the ***joint probabilities***, and the numbers in the right and bottom margins of the table are ***marginal probabilities***. The table has exactly the same information content as either probability tree.

	e_1 = D	e_2 = $\bar{D}$	
"A"	.09	.18	.27=P("A")
"$\bar{A}$"	.01	.72	.73=P("$\bar{A}$")
	.10 = P(D)	.90 = P($\bar{D}$)	

If we can obtain the seismic information before making a drilling decision, then the probability of making a field discovery can be revised based on the additional information. Bayes' rule may be used, or one can inspect the probability table to obtain the revised probabilities. For example, if the seismic survey indicates an anomaly (given "A"), then only the top row of the table remains relevant. The

0.09 and 018 probabilities are in the right proportion for D and $\bar{D}$, however they do not sum to one. ***Normalize*** these joint probabilities to obtain the revised probabilities. The probability of discovery given anomaly, P(D|"A"), is .09/.27. The expression is read 'the probability of D given "A".' Similarly, we can extract the probability of no

discovery given anomaly, P($\bar{D}$ |"A") = .18/.27.

Instead of setting up and inspecting the joint probability table, the analyst may want to represent the calculations with the Bayes' rule. A revised probability for event e_i, given event A can be determined using Bayes' Rule:

$$P(e_i|A) = \frac{P(A|e_i)P(e_i)}{\sum_{\text{all } j} P(A|e_j)P(e_j)}$$

Note that the denominator in the equation is equal to P(A), and the numerator is $P(A \cdot e_i)$. Bayes' rule is just a generalization of the *multiplication rule*.

For our example,

$$P(D|A) = \frac{P(A|D)\,P(D)}{P(A|D)\,P(D) + P(A|\bar{D})\,P(\bar{D})}$$

$$= \frac{.9(.1)}{.9(.1) + .2(.9)} = \frac{.09}{.09 + .18} = \frac{1}{3}$$

$$P(D|\bar{A}) = \frac{P(A|D)\,P(D)}{P(\bar{A}|D)\,P(D) + P(\bar{A}|\bar{D})\,P(\bar{D})}$$

$$= \frac{.1(.1)}{.1(.1) + .8(.9)} = \frac{.01}{.01 + .72} = 0.0137$$

The other branch's probabilities at each Discovery node are complementary and are determined by subtracting from 1.