

II. THE MONTY HALL AND OTHER PROBABILISTIC PROBLEMS

Probability considerations are everywhere — selecting apparel when one is going out, chances for rain, catching your bus, getting caught in traffic. Perhaps you learn when you get to work that your boss has handed you a pink slip. What are your chances for successfully stopping smoking? What is the probability that your house will have been robbed before you return home? At dinnertime will your wife tell you she's pregnant? Will the police call to tell you that your stolen car has been recovered — but with the wheels missing? Will your income tax return be audited? Will your brother die of bone cancer? What is the chance you will be hit by lightning or win \$11 million in the Publisher's Clearing House lottery? You might encounter these and many other probabilistic situations on an almost daily basis. Problems involving the calculation or estimation of probabilities arise in every academic area, every workplace, and during your leisure hours as well. Given below are four introductory problems that involve notions of probability. See what you can do with them. Perhaps you can make significant headway with one or more. Or perhaps you will not be able to get started at all. After making a serious effort to handle these problems you can then remedy your deficiencies by reading the background materials that follow. Upon returning to these same introductory problems you should then fare much better!

In this problem solving program we will be dealing with three quite distinct kinds of probability. The first of these is "objective" probability, so-called because of what it's not, and it's not subjective. It is standard text book probability based upon the idea of equi-probable events. Example settings include the tossing of a coin, the throw of a pair of dice, drawing from a deck of cards, or blindly selecting colored beads from an urn. It is only because there are events with equal probability that we know that in the single toss of a coin the probability that it will come up heads is the same as for it coming up tails. For this same reason we know that every possible outcome in the throw of a pair of dice has a probability of $1/36$. This makes the probability for a single outcome, say of two 6's, one in thirty-six because there is only one way to obtain this particular outcome. Similarly, the probability for drawing a Queen of Spades from a full pack of playing cards is $1/52$ or for drawing any one of the four Queens in the deck $4/52$. The second kind of probability may be called "sample probability." If in 1000 individuals 200 become afflicted with stomach ulcers after having taken ibuprofen for a month, we say that the "probability" for getting a stomach ulcer if one takes ibuprofen for a month is $200/1000$, or 20%. Since there are considerably more individuals than 1000 taking ibuprofen, the group of 1000 individuals is but a sample. The probability value of 0.20 is likely to be fairly close to the probability figure that would be appropriate for all individuals taking ibuprofen for a month, and so we use that figure. If the size of the sample is relatively small, however, a probability figure based upon that number would of course be highly suspect. The third kind of probability is called "subjective," or personal, probability. For example, it can be used to express the probability that the earliest Americans came from Asia across a land bridge between the continents where now there is the Bering Straits. Suppose we set this probability to be 0.99. It is of course ridiculous to say that in 100 opportunities for these early Asians to cross such a bridge to America, in 99 cases they seized the opportunity and on one occasion they did not, making the probability for crossing 0.99. Other examples of subjective probability include the probability for a horse to win a race, the probability for an incumbent to win a second term in office, and the probability for the stock market to crash this month. A more complete investigation of subjective probability will be made in a later chapter. (vii).

Introductory Problems

THE MONTY HALL PROBLEM. This problem was presented by Marilyn vos Savant in Parade Magazine. It involves the game show Let's Make A Deal hosted by Monty Hall. Monty gives a game contestant the opportunity to choose one of three closed doors. Behind one of these doors is a new automobile but behind each of the other two there is a tethered goat. The contestant gets to keep whatever is behind the door he or she selects. Let us call these doors door A, door B, and door C. With considerable "help" from the audience the contestant selects one of the doors. Suppose it is door A. Before opening this door, however, Monty opens one of the other two doors, say it is door B, to reveal a goat. At this point Monty gives the contestant an opportunity to change the original selection of door A to that of door C.

The question is a simple one. Should the contestant switch to door C or remain with the original choice of door A, or doesn't it make any difference? Of course it is not certain which of these two doors conceals the new car. It is a matter of probability. Is the probability greater to find the auto behind door A or behind door C or are the probabilities equal? It is important to recognize that although the contestant doesn't know where the new car is, Monty does.

BLUE AND GREEN TAXICABS. This problem is described by the psychologists Kahneman and Tversky in *Judgment under uncertainty: Heuristics and biases*. A cab was involved at night in a hit and run accident. In this town there are only two cab companies, the Blue with 85% of the cabs and the Green company with the remaining 15%. As it happened there was a witness to the accident who identified the hit and run cab as Blue. The case came to trial. An important part of the testimony was a test of the reliability of the eyewitness to identify a cab as being either Blue or Green under the same circumstances as those that existed the night of the accident. The result of these tests was that the witness correctly identified a blue cab as blue 60% of the time and a green cab as green 80% of the time. On the basis of the claim of the eyewitness and the tests conducted by the court, what is the probability that the cab involved in the accident was in fact Blue as the witness had claimed?

MAMMARY CANCER. The following is a case adapted from one described by David Eddy. A physician has a patient with a breast mass. On the basis of his rather extensive experience with such patients of the same age, family history, and symptoms he is somewhat suspicious that the breast mass is cancerous. He assigns a subjective probability of ten percent to the possibility that this patient has breast cancer. The doctor orders that a mammogram be taken. In the radiologist's opinion the lesion is malignant. At this point the physician does some library research. The best evidence he could find regarding the accuracy of mammography is a study with the following results: mammography correctly diagnosed 79.2 percent of 475 lesions as malignant and correctly diagnosed 90.4 percent of 1,105 lesions as benign. On the basis of this information, taking account of the patient's mammogram, the physician revises his opinion. What probability should he now assign to the possibility of cancer?

CASE OF THE RODEO GATECRASHERS. A hypothetical case in civil law is described by Allen in *Probability and Inference in the Law of Evidence*. A thousand people are seated in the stands at a rodeo. However, only 499 tickets have been sold. Since no ticket stubs were given to any of the ticket purchasers, there is no way of telling which of the 1000 individuals in the stands bought their tickets and who "climbed over the fence." The rodeo organizers single out a particular individual, call him A, as a test case and charge him with having climbed over the fence. It would appear that the rodeo organizers are entitled to judgment against A for the admission money since, in a civil case, one needs only to show guilt by a "preponderance" of the evidence. Clearly, the probability that A is guilty is 501/1000, a value exceeding 50 percent and therefore sufficient to make the case for the rodeo organizers. The absurdity of this situation is evident since, if the price of admission can be claimed against A it can also be claimed against the other 999 people in the stands. The management would then stand to collect the price of 1000 admissions in addition to the 499 that were already paid for. Something is not quite right here, but what is it?

Background

BASIC PROBABILITY RULES. In our work we will be dealing with only two propositions at a time. Call these proposition or statement A and proposition or statement B. Understand that A and B are merely symbols that can stand for almost any kind of affirmative statement. A could be *it is raining* and

B could represent it is cold. Either A or B could say: *this patient has cancer, or the test for AIDS was negative*. We will restrict ourselves for the time being to statements that are either true or false, nothing in-between. $P(A) = 0.6$ means only that the fraction of times A is true is 0.6. It is recognized that this condition is quite unrealistic in the real world where propositions such as *our foreign policy in China will be successful* lurk everywhere. There are three important mathematical rules. These are:

$$P(A) + P(\bar{A}) = 1$$

$$0 \leq P(A) \leq 1$$

$$P(A) = P(A \cdot B) + P(A \cdot \bar{B})$$

where $P(A)$ expresses the probability that proposition A is false and $P(A \cdot B)$ is read as the probability of A and B, i.e., that both are true. The first equation states that a proposition such as A must be either true or false since the probability that it is true when added to the probability that it is false must be one, the probability that it is one or the other. The second states that probability values are never negative and never greater than 1. $P(A) = 0$ and $P(A) = 1$ express, respectively, the certainty that A is false and the certainty that it is true. The third expression tells us that $P(A)$ is the sum of the probabilities of A "anded" with each of B and \bar{B} , i.e., that $P(A)$ is the probability that A is true and B is true added to the probability that A is true and B is false.

EQUI-PROBABLE EVENTS. In the roll of a pair of fair dice each possible outcome is as probable as every other outcome. One must, however, think of each die as being red and the other as green. That is, imagine one die as being red and the other as green. In this way a roll that yields a red 3 and a green 4 is distinct from one that yields a red 4 and a green 3. With 36 possible outcomes the probability P for each outcome is $1/36$. If one wishes to know the probability for rolling a seven, one need only count the "ways" this can occur: 1 and 6, 2 and 5, 3 and 4, 4 and 3, 5 and 2, and finally, 6 and 1. With six ways, the probability $P(7)$ becomes $6 \times 1/36$, or $1/6$. Another equi-probable situation is found when drawing blindly from an urn which contains colored beads. Suppose an urn contains 7 red beads and 14 white beads. There are then 21 possible draws of a single bead, each of which is as likely as any other. $P(\text{red})$ is therefore $7/21$, or $1/3$, and $P(\text{white})$ is $14/21 = 2/3$. An ordinary deck of playing cards contains 13 Spades, 13 Hearts, 13 Diamonds, and 13 Clubs making 52 cards in all. In each suit the cards are 2 through 10 followed by a Jack, Queen, King, and Ace. With the deck well shuffled and face down the probability for drawing the Queen of Spades is $1/52$ since each card is as likely to be drawn as any other. The probability for drawing any Queen is $4/52$ and for drawing a Heart is $13/52 = 1/4$. Finally, for each toss of a fair coin the probability for tossing a head is the same as that for tossing a tail, viz., $1/2$.

CONDITIONAL PROBABILITY. A conditional probability is written in the form $P(A|B)$ and read as the probability of A on condition that B is true. The short vertical line between the A and the B represents the language on condition that, or given that. For example, the probability for drawing an Ace from a deck of cards on condition that another Ace has already been drawn from the deck can be written $P(A_2|A_1)$ where the subscripts indicate the second draw and the first respectively. Whether this probability equals $P(A_2)$ depends on whether the first Ace is returned to the pack before making the second draw. In drawing colored beads from an urn the probability for drawing a particular color depends on whether prior draws were or were not returned to the urn after each draw. In tossing a coin time after time the situation is similar to that for repeated throws of a pair of dice. The probability for tossing a head $P(H)$ is $1/2$ no matter what the previous throws might have been. We say that each throw is *independent* of all other throws. In the case of the woman whose physician thinks there is only one chance in ten that the lump in her breast is malignant, the quantity that is sought is the probability that the woman has cancer on condition that she has tested positive. This can be abbreviated as $P(C|+)$. For conditional probabilities the first rule above becomes $P(A|B) + P(A|\bar{B}) = 1$, and it is not true that $P(A|B) + P(A|\bar{B}) = 1$.

AND'S and OR'S. Time and time again we will want to know the probability of one statement, or proposition, call it A, and a second proposition or statement, call it B. This is often written $P(A \cdot B)$, or

sometimes, without the period inbetween the A and the B. This is called the conjunctive probability of A and B (recall that the word *and* is a conjunction). The probability of A or B is written $P(A \vee B)$. This *or* means that either A is true or B is true or both are true. It is therefore the inclusive *or* as opposed to the exclusive *or*. The exclusive *or* means that either A is true or B is true but not both. There are rules for calculating both $P(A \cdot B)$ and $P(A \vee B)$, a rule for each that applies in general, and a second rule for each that is used if a certain condition is met. This makes four rules in all. These are important!

and rule:
$$P(A \cdot B) = P(A)P(B|A) \\ = P(B)P(A|B) \dots \text{both are valid expressions in general}$$

if A and B are independent: $P(A \cdot B) = P(A)P(B)$ since in this case $P(B|A) = P(B)$
and $P(A|B) = P(A)$

or rule:
$$P(A \vee B) = P(A) + P(B) - P(A \cdot B) \dots \text{valid in general}$$

if A and B are mutually exclusive:

$$P(A \vee B) = P(A) + P(B) \dots \text{since here } P(A \cdot B) = 0$$

In the above *independence* as applied to the *and* rule means that proposition A is in no way related to proposition B. For the *or* rule, when two propositions are mutually exclusive it means that they cannot occur together. Recall the basic probability rule

$$P(A) = P(A \cdot B) + P(A \cdot \bar{B})$$

Using the *and* rule this can now be rewritten as

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$$

These rules will have more meaning when we see them play a role in several examples. Consider a deck of playing cards. Recall that Spades and Clubs are both colored black while Hearts and Diamonds are both red. Let us draw one card from the deck and then, without replacing this card in the deck, draw a second card. The probability that both of these cards are red is:

$$P(R_1 \cdot R_2) = P(R_1)P(R_2|R_1) = (1/2)(25/51)$$

since after drawing the first card there are only 25 red cards left in a pack of 51 cards. If however the first card drawn is replaced before the second draw the two draws are independent of each other so that

$$P(R_1 \cdot R_2) = P(R_1)P(R_2) = (1/2)(1/2) \dots \text{independence}$$

To illustrate the *or* rule consider a deck of cards from which a single card is taken. The question is this: What is the probability for drawing a Queen or a Heart? Since these are not mutually exclusive possibilities, we use the *or* rule to obtain:

$$P(Q \vee H) = P(Q) + P(H) - P(Q \cdot H) \\ = 1/13 + 1/4 - (1/13)(1/4) = 4/52 + 13/52 - 1/52 \\ = 16/52 = 4/13$$

This may appear as an odd expression for the probability for selecting in a single draw a Queen or a Heart (this *or* is the inclusive *or*). What's odd is the subtracting off of the quantity $P(Q \cdot H)$.

Bayes' equation is written in its simplest possible form below. Also noted below are the various roles it can play.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \dots \text{turns } P(B|A) \text{ into its converse } P(A|B)$$

verosimilitud de A, dado B

probabilidad de A, dado B

... converts a prior value $P(A)$ into a posterior updated value $P(A|B)$ having learned that B is true

$$P(B) = P(B \cdot A) + P(B \cdot \bar{A})$$

$$= P(A \cdot B) + P(\bar{A} \cdot B)$$

$$= P(A)P(B|A) + P(\bar{A})P(B|\bar{A})$$

... is a rewritten form of $P(A|B)P(B) = P(B|A)P(A)$, each side of which expresses $P(A \cdot B)$

The above expression is a satisfactory one for finding $P(A|B)$ provided that the three quantities on the right side of the equation, $P(B|A)$, $P(A)$, and $P(B)$ are all known. But in this chapter $P(B)$ is *not* known and $P(B|\bar{A})$ is known. We therefore in these cases use, not $P(B)$ itself, but what it equals which is $P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$. Substituting this expression into the denominator of the expression above gives:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} \dots \text{Bayes' equation when } P(B|A), P(A), \text{ and } P(B|\bar{A}) \text{ are all known}$$

Including the version of Bayes' equation written above, there are in all a total of four such expressions. It is useful to be able to write all of these correctly. The remaining three are:

$P(A|\bar{B}) = \dots$ same as the above with \bar{B} substituted for B everywhere

$P(B|A) = \dots$ same as the first expression with A and B interchanged everywhere

$P(B|\bar{A}) = \dots$ same as expression directly above with A replaced by \bar{A} everywhere

It may be easier to remember all these expressions when one concentrates on the *pattern* exhibited by all the symbols. This pattern can perhaps be better seen if it is written in Greek:

$$P(\Psi|\Phi) = \frac{P(\Phi|\Psi)P(\Psi)}{P(\Phi|\Psi)P(\Psi) + P(\Phi|\bar{\Psi})P(\bar{\Psi})} \dots \text{Bayes' equation relating propositions } \Psi \text{ and } \Phi$$

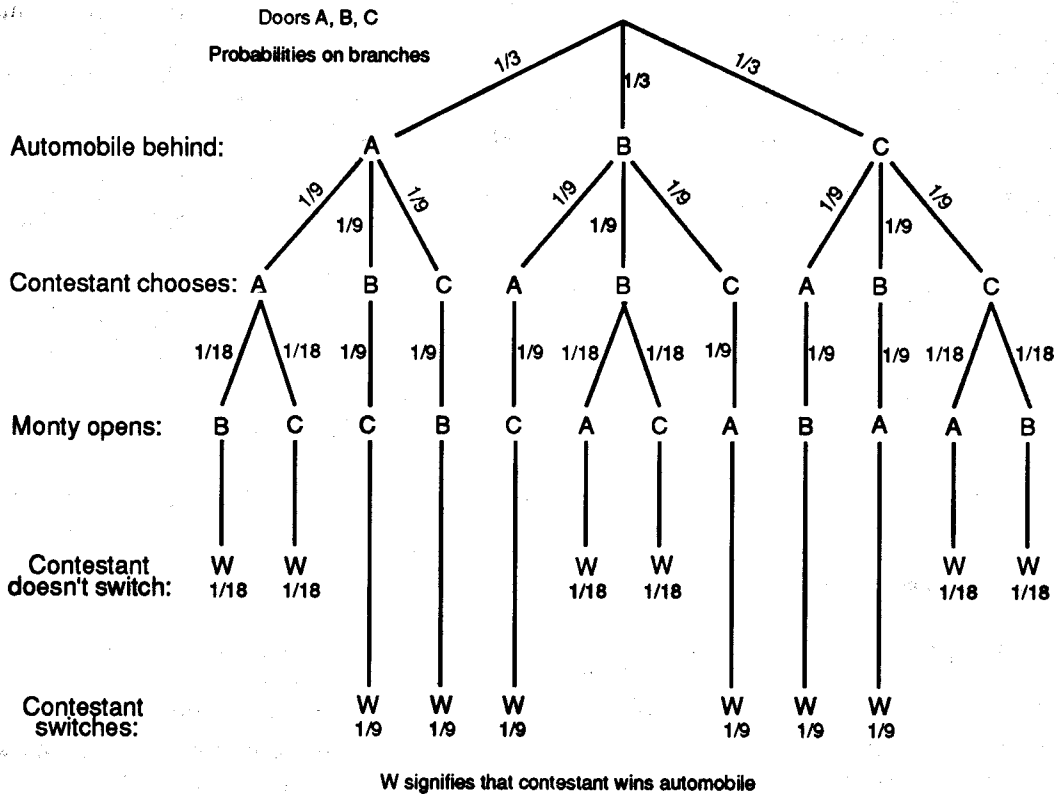
Here, Ψ can represent the proposition C that the patient has cancer, as it does in the MAMMALARY CANCER problem, and then Φ represents the proposition + indicating a positive result from a mammogram. Then, of course $\bar{\Psi}$ represents the proposition that the patient does not have cancer and $\bar{\Phi}$ the proposition that the mammogram gives a negative result.

But what about the conditionals $P(\bar{A}|B)$, $P(\bar{A}|\bar{B})$, $P(\bar{B}|A)$, and $P(\bar{B}|\bar{A})$? Each of these is directly obtainable by taking, in turn, 1 minus $P(A|B)$, $P(A|\bar{B})$, $P(B|A)$, and $P(B|\bar{A})$. These are relationships that are related to $P(A) + P(\bar{A}) = 1$ and $P(B) + P(\bar{B}) = 1$, differing only in the fact that they are conditionals.

In setting up Bayes' equation in the analysis of a problem, one needs to be sensitive to the various ways a conditional probability can be expressed. For instance, in the MAMMALARY CANCER problem, one piece of data that is given is that $P(+|C) = 0.792$. Symbolically this is unambiguous. However, it could have been stated as *the probability for a positive result on condition that the patient has mammary cancer*, or as, *the probability for a positive result given that the patient has mammary cancer*, or as, *the probability for a positive result if the patient has mammary cancer*. Also, things could be turned around by stating that *if the patient has cancer the probability for a positive result is 0.792*, or as, *given that the patient has cancer the probability for a positive result is 0.792*, or as, *on condition that the patient has cancer the probability that the test will be positive is 0.792*. Finally, one could say that *the probability is 0.792 that a patient with cancer will receive a positive test*.

Introductory Problem Solutions

THE MONTY HALL PROBLEM. DRAW TREE. A decision tree for the Monty Hall problem is given below. It is assumed that the producers of the television show select at random the door behind which the automobile is to be placed before the program begins. Thus the probability for it to be behind any particular door is $1/3$. It is assumed also that contestants in general have no particular preference for any one of the letters A, B, or C. This makes the initial selection of a particular door by a contestant as probable as the selection of any other door, each having probability $1/3$. There are at this time nine possibilities, each one as probable as the others. The probability that the car is behind a given door combined (the "and" rule; recall that "ands" multiply) with the probability that the game contestant will select a particular door makes the probability of each possibility $1/9$, as shown.



Conclusion: $P(W \mid \text{contestant doesn't switch}) = 1/3$

$P(W \mid \text{contestant switches}) = 2/3$

But now comes a deviation in the calculation of probabilities. After the contestant has selected one of the three doors it is now time for Monty to open one of the doors to show the contestant that the automobile isn't there, only a goat.

RECOGNIZE SITUATION. What makes this situation different is that Monty *knows* where the automobile is. We know that he isn't going to open the door in front of the automobile. If the contestant has happened to select the door leading to the automobile Monty has two choices. He can open either of the other two doors to show the contestant that the automobile isn't there. But if the contestant has selected the wrong door, then Monty has only one choice. He isn't going to open the door the contestant has selected and he isn't going to open the door in front of the automobile. This leaves him with only one choice for the door he will open before the contestant must decide whether he is going to stick with his original door selection. When Monty has two choices we assume his selection of one of these two is

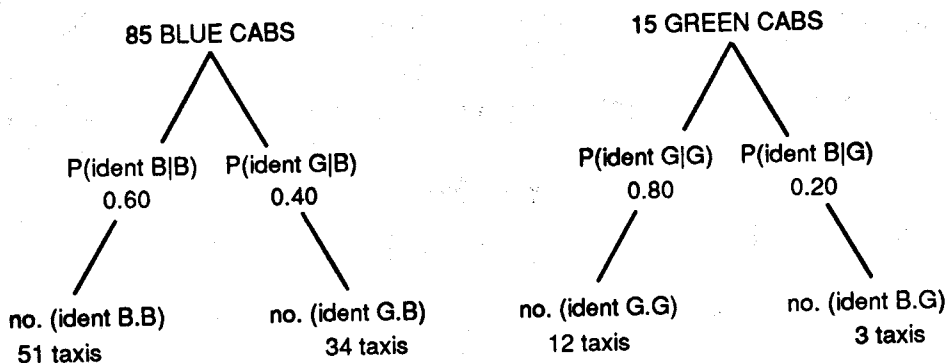
random and each possibility has an associated probability of 1/18. When Monty has but one choice the probability for arriving at this juncture is 1/9. This is the feature that makes the probability for the contestant to win the car greater if the contestant switches her choice. As can be seen from the decision tree, if she "sticks" her chances are 6/18 = 1/3 while if she "switches" her probability to win is 6/9 = 2/3.

BLUE AND GREEN TAXICABS. IDENT PROB TYPE. In this problem there are but two statements. First, the hit-and-run-cab is blue or it is not blue (green). Second, the eyewitness either correctly identifies this cab as blue or incorrectly as green. This assessment leads to two distinct methods for attacking this problem.

IDENT FIRST METHOD. Suppose the test of the reliability of the eyewitness consisted of her identification of 100 cabs, 85 of which were blue and 15 of which were green. From the given information one can then determine how many of these fell into each of the following categories:

number (ident B.B) number (ident G.G)
 number (ident G.B) number (ident B.G)

A "tree" will organize the data which is shown in the diagram below:



SOLVE. Since the eyewitness identified the hit-and-run cab as blue, the probability that the cab actually was blue is the quantity $P(B|\text{ident B})$ which can be determined numerically from the conjunctive probabilities above as follows:

$$\frac{\#(\text{ident B} \wedge B)}{\#(\text{ident B})} = P(B|\text{ident B}) = \frac{51}{51 + 3} = 0.94$$

Note that this method makes no explicit use of Bayes' Equation and could have been used to solve this introductory problem before the background information was known.

IDENT ALTERNATIVE METHOD. Bayes' Equation is ideally suited for the analysis of this problem. This method is more abstract than the method used above but is more generally applicable. We use it in the following form where the denominator is $P(B)$.

$$P(B|\text{ident B}) = \frac{P(\text{ident B}|B) P(B)}{P(\text{ident B}|B) P(B) + P(\text{ident B}|G) P(G)}$$

Numerically, $P(B|\text{ident B}) = 0.94$

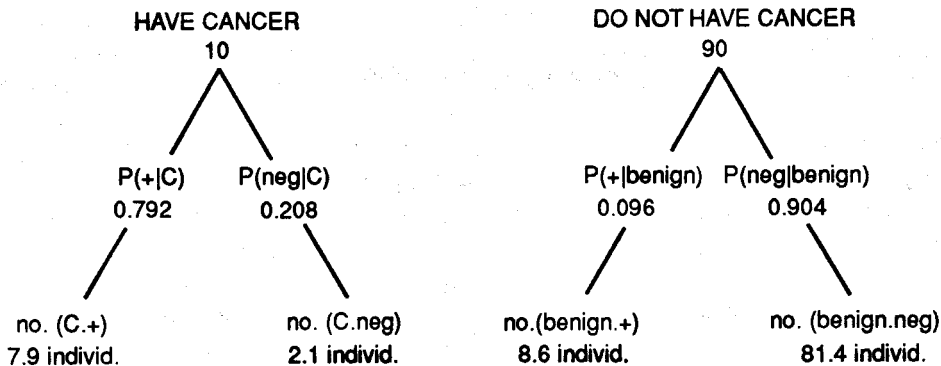
MAMMARY CANCER. IDENT PROBLEM TYPE. This problem seems to be similar to the problem of the BLUE AND GREEN TAXICABS because here, as there, there are two propositions: the individual either has cancer (C) or not (benign), and the mammogram is either positive (+) or negative (neg), i.e., is not positive. We expect there to be two methods for finding a solution, just as there were for the taxicabs, the first involving a tree structure and the second using Bayes' Equation.

IDENT GIVENS. As with most problems, one of the first steps is to summarize the information that is given. In this case:

- $P(C)$ = prior probability that the patient has cancer = 0.10
- then $P(\text{benign})$ = prior probability that the breast mass is benign = 0.90
- also $P(+|C)$ = probability for test to be + given that patient has cancer = 0.792
- $P(+|\text{benign})$ = probability for + test if breast mass is benign = 1 - probability for a correct negative test if breast mass is benign = $1 - P(\text{neg}|\text{benign}) = 1 - 0.904 = 0.096$

The structure of the problem is clearly the same as that for the problem of the BLUE AND GREEN TAXIS. The two methods for attacking the problem will therefore be the same.

FIRST METHOD. Suppose there are 100 individuals with the same diagnosis as that for the patient in this problem.



WANTED: The desired quantity is $P(C|+)$, the probability that the patient has cancer now that her mammogram is determined to be positive.

$$P(C|+) = \frac{7.9}{7.9 + 8.6} = 0.48$$

ALTERNATIVE METHOD. The second method involves the use of Bayes' Equation.

SOLVE. The givens are the quantities that go into Bayes' Equation written as follows:

$$P(C|+) = \frac{P(+|C) P(C)}{P(+|C) P(C) + P(+|\text{benign}) P(\text{benign})}$$

Numerically, $P(C|+) = 0.48$

CASE OF THE RODEO GATECRASHERS. RESTATE PROBLEM. Recall that there were 1000 people in the stands at the rodeo but only 499 paid for admission. The obvious conclusion is that 501 individuals were "gatecrashers." It seems absurd that the rodeo management might select one of the thousand