PROBABILITY

Documents prepared for use in course B01.1305,
New York University, Stern School of Business

Types of probability page 3

The catchall term *probability* refers to several distinct ideas. This discusses also the axioms of probability and the notion of a fair bet.

Arbitrage of probability page 5

Personal styles influence the types of bets that people make. This shows also what happens when two people disagree about the probabilities.

Conditional probability examples page 7

Here are some simple illustrations of conditional probability.

Expected value examples page 10

The expected value concept is illustrated here for some simple games of chance.

Probability trees to find conditional probabilities page 13

This shows a schematic that's useful for finding conditional probabilities in some contexts. The ideas are also illustrated through a two-by-two table, which might be easier.

Efron's dice page 16

Probabilities can surprise you.

Let's Make a Deal page 18

There's an interesting probability controversy related to this old television show, and it's become part of the cultural heritage of probability.

The Prosecutor's Fallacy page 21

Juries must make yes-or-no decisions based on probabilities related to the evidence. This shows a very common abuse of the probability notions.

Edit date 8 JAN 2007

Cover photo: Rabbit, San Diego, California, 2005

⋞⋞⋞⋞⋞⋞TYPES OF PROBABILITY⋟⋟⋟⋟⋟⋟

There are some primitive facts about probability that are understood easily:

For any event E, $0 \le P[\, E\,] \le 1$.

$P[\, E\,] = 0$ means that E is impossible.

$P[\, E\,] = 1$ means that E is certain.

If E' denotes "not-E," then $P[\, E'\,] = 1 - P[\, E\,]$.
Equivalently, $P[\, E\,] + P[\, E'\,] = 1$.
You'll also see $\overline{E}$ or E^c or $\neg E$ as other symbols for "not-E."

If E and F are events that cannot happen together (such as getting "1" and getting "4" on the same roll of a die), then $P[\, E \cup F\,] = P[\, E\,] + P[\, F\,]$. We use " E ∪ F " to stand for "E or F."

Where do probabilities come from?

1. Some probabilities are based on physical properties of the equipment. The standard games of chance using coins, dice, roulette wheels are such situations. There is no argument that P[heads] = $\frac{1}{2}$ for one flip of a coin or P["two dots"] = $\frac{1}{6}$ for a single role of one die.

2. Some probabilities are based on the long-range frequentist interpretation. Consider an event E, perhaps representing the event that a flipped thumb tack will land point-up. $P[\, E\,]$ is not known, but we believe that we can estimate it with a long enough experiment. Our estimate for $P[\, E\,]$ is $\frac{\text{Number of times } E \text{ occurs in } n \text{ trials}}{n}$. That this estimate converges to $P[\, E\,]$ is the "law of large numbers," which is more commonly known as the "law of averages." The law of averages applies also to situations noted in type 1. Since the probabilities are known in those situations, there is no need to run the experiments to estimate the probabilities (although there is some interest in assessing the rate of convergence).

3. Some probabilities are based on mathematical models. As an example, consider the problem of assessing P[rain tomorrow]. You can hear such numbers in weather forecasts. These cannot be based on the law of averages, since it is impossible to replicate the weather conditions even once, let alone infinitely many times. The actual technique involves substituting various meteorological measurements into a mathematical formula to produce the probability. The mathematical formula is developed over a period of time, and the objective is to produce a bet-either-side number. The meaning of "bet either side" is explained below.

4. Some probabilities are subjective. Betting on sports events requires forming subjective probabilities. Note that *X* 's subjective probability for P[*A*] is that number which produces for him a bet-either-side situation.

5. Some probabilities are the result of a market arbitrage of subjective probabilities. If many people are betting on a sporting event, then their subjective probabilities are subjected to a negotiation process. This process eventually produces a betting line. At pari-mutual race tracks, the arbitrage process is computerized.

We should mention here the Bayesian treatment of probabilities. The Bayesian theology is a bit too complicated to summarize completely here, but its centerpiece concept is that subjective probabilities have all the dignity and practical usefulness of the other varieties of probability. Further, the Bayesians are willing to place a probability structure on quantities which are fixed but are unknown. In point (2), for example, a Bayesian is quite willing to express probabilities about a thumb tack's landing with its point upward *even without performing a single trial*; the probability statements are revised as the experiment proceeds.

Finally, let's note what we mean by a bet-either-side situation. Suppose that it is known by all that P[*E*] = 0.37. Here is how we formulate a bet:

> One side puts up the amount \$37 and wins \$100 if *E* happens.
>
> The other side puts up the amount \$63 and wins \$100 if E' happens.

This is also called a *fair bet* and it has the property that a lover of gambling would be willing to bet on either side. That is, it has to appear to the gambler that neither side has an advantage over the other.

An example of a fair bet would be betting even money, heads versus tails, on a coin flip. The gambler really doesn't care whether he's betting on heads or on tails.

> The \$100 in this example is arbitrary. It's just a large enough value to make the game interesting. These people would probably not be involved in a game wagering 37¢ against 63¢, and they would probably be scared out of a game wagering \$37,000 against \$63,000.

⋖⋖⋖⋖⋖⋖ARBITRAGE OF PROBABILITY⋗⋗⋗⋗⋗⋗

In this document we will consider the role that subjective probability plays in gambling decisions. It will become apparent that probabilities are thus subject to a form of arbitrage.

One can divide gamblers into categories according to the states of wisdom under which they are willing to operate.

> Some gamblers bet for the thrill and will involve themselves in situations that are highly unfavorable. Casino gamblers who play games like roulette are in this category. The game methodically and with grinding regularity separates them from their money, but they play anyway. These people tend to have a poor understanding of probability.
>
> Unfavorable bets for large stakes can be quite rational. Life insurance and catastrophic health insurance are unfavorable bets for the policyholders, but such insurance is certainly a wise purchase.
>
> The purchase of unfavorable lottery tickets (for high prizes) can be intellectually justified. After all, many people enjoy the thrill of state lotteries in which the largest prizes can be over $100,000,000.
>
> Lottery games for small prizes are silly bets. So are insurance policies that cover losses that one could easily pay out-of-pocket. Small-stakes state lottery games have been described as “a tax on people who don’t understand probability.” (The quote was picked up on the Internet, but the author is not known.)
>
> There are also gamblers who like to play complicated faddish games in which they are likely to outsmart novices. The craze for Internet poker games (like Texas hold-‘em) illustrates this point.
>
> Some gamblers will insist that the games be at least close to favorable. These people tend to play blackjack and craps.
>
> Some gamblers will always rule out unfavorable bets. They will not play casino games, but they are willing to take even bets, such as putting even money on the result of a coin flip.
>
> Some gamblers insist on favorable probabilities. Many bettors at racing tracks are in this category (though they are frequently mistaken about the probabilities).

Let’s note how a fair bet works out in a gambling situation. If it is agreed by X and Y (gamblers in the final category above) that $P[E] = c$ and $P[E'] = 1 - c$, then there are indifference bets (even bets, fair bets, bet-either-side).

> X is willing to bet amount $c\ M$ for event E to occur.
> Y is willing to bet amount $(1 - c)M$ for event E' to occur.

The game will then be played, and either E or E' will happen. If E happens, then X wins the total amount M. If E' happens, then Y wins the total amount M. Each of X and Y regard this as a fair bet.

> ASIDE: M is big enough to make the game interesting (but not so big that the game is scary). Perhaps $c = 0.4$ and $M = \$10$, so that the bets are \$4 and \$6.

Each gambler puts up an amount proportional to the probability of the event he is choosing. That is, X 's bet is proportional to P[E] and Y 's bet is proportional to P[E'].

When P[E] and P[E'] are common knowledge, and both gamblers are willing to take fair bets, then the two players would be willing to exchange positions.

Now let's consider a case in which the gamblers disagree about the probabilities. We'll do this with a specific example. Suppose that teams A and B are about to play a basketball game. Let P[A] denote the probability that team A will win. Let P[B] = 1 - P[A] denote the probability that team B will win. Since this is nothing like a coin flip, it is easy to imagine that there may be disagreements about the probabilities.

> Gambler X thinks that P[A] = 0.40 and P[B] = 0.60. Gambler X is willing to bet (\$40, A) against anyone else who will do (\$60, B). He is also willing to take the other position, betting (\$60, B) against anyone else taking (\$40, A).

> Gambler Y thinks that P[A] = 0.30 and P[B] = 0.70. Gambler Y is willing to bet (\$30, A) against anyone else who will do (\$70, B). He is also willing to take the other position, betting (\$70, B) against anyone else taking (\$30,A).

Now suppose that X and Y happen to meet. Because there is a difference of opinion here, a bet is certainly going to happen. The bet will be

> Gambler X takes (\$35, A).
> Gambler Y takes (\$65, B).
>> Actually, the \$35:\$65 split could end up as \$34:\$66 or \$38:\$62. The numbers depend on the negotiating skills of X and Y.

Gambler X regards the bet as a bargain. He gets team A, but he wagers *less* than \$40 with the change of winning *more* than \$60.

Gambler Y regards this as a bargain. He gets team B, but he wager *less* than \$70 with the chance of winning *more* than \$30.

Note that the gamblers are *not* willing to exchange positions!

Situations in which there are disagreements about probabilities are those which lead to bets. Many financial transactions are based on such disagreements.

⋖⋖⋖⋖⋖⋖CONDITIONAL PROBABILITY EXAMPLES⋗⋗⋗⋗⋗⋗

This document gives a number of examples of probability problems, including conditional probability.

EXAMPLE: Suppose that three slips of paper have the names *a*, *b*, *c*. Suppose that these are given at random to people with names *A*, *B*, *C*. What is the probability that exactly one person gets the paper matching his or her name?

SOLUTION: There are six possible equally-likely outcomes to this situation:

Paper received by		
A	*B*	*C*
a	*b*	*c*
a	*c*	*b*
b	*a*	*c*

Paper received by		
A	*B*	*C*
b	*c*	*a*
c	*a*	*b*
c	*b*	*a*

Each situation has probability $\frac{1}{6}$. There are exactly three situations (those shaded) in which exactly one person gets the matching paper. The probability is then $\frac{3}{6} = \frac{1}{2}$.

EXAMPLE: The probability that a train leaves on time is 0.85. The probability that it leaves on time and arrives on time is 0.60. If it leaves on time, what is the probability that it will arrive on time?

SOLUTION: Let's make a diagram which shows the possibilities. It may help to imagine a fictitious set of 1,000 trips. Of these, 850 will leave on time. Similarly, 600 trips will leave on time and also arrive on time. The state of our knowledge is this:

	Arrive on time	Arrive late	Total
Leave on time	600		850
Leave late			150
Total			1,000

We can certainly place the value 250 in the box (Leave on time, Arrive late). The bottom row cannot be determined from the information given. We know only this:

	Arrive on time	Arrive late	Total
Leave on time	600	250	850
Leave late			150
Total			1,000

The question now asks.....What is the probability that a train which leaves on time will arrive on time? We have 850 fictitious trains leaving on time, and 600 of them arrived on time. Our probability is thus $\frac{600}{850} \approx 0.7059 \approx 0.71$.

You can solve this in symbols as well. Let A be the event "leaves on time" and let B be the event "arrives on time." We are given P(A) = 0.85 and P($A \cap B$) = P(A and B) = 0.60. Then $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.60}{0.85} \approx 0.7059$.

The next examples ask for a literal interpretation of conditional probabilities.

EXAMPLE: Flip a coin. If the coin comes up heads, select one ball from jar A. If the coin comes up tails, select one ball from jar B. Jar A contains 8 red and 2 green balls. Jar B contains 15 red and 30 green balls. Find P[red | heads] and P[red | tails].

SOLUTION: P[red | heads] should be interpreted as the probability of getting a red ball, *given* that the coin flip was a head. This means that the ball is to be selected from jar A, and you can rephrase the question as P[red | A]. Now, P[red | A] = $\frac{8}{8+2} = 0.8$.

Similarly, P[red | tails] = P[red | B] = $\frac{15}{15+30} = \frac{1}{3} \approx 0.3333$.

EXAMPLE: You have a standard deck of 52 cards. Select one card. Find P[heart | ace].

SOLUTION: You are given that the selected card is an ace. Since there are four aces, and since just one of these four aces is also a heart, you should find P[♥ | ace] = $\frac{1}{4}$ = 0.25.

You can formally apply the definition of conditional probability to write

$$P(♥ | ace) = \frac{P[♥ \cap ace]}{P[ace]} = \frac{P[A♥]}{P[A]} = \frac{\frac{1}{52}}{\frac{4}{52}} = \frac{1}{4}$$

Let's repeat the previous example, but do it with a defective deck of cards.

EXAMPLE: You have a deck of 51 cards, which happens to be a standard deck missing the eight of diamonds. If you select one card from this deck, find P(♥ | ace).

SOLUTION: $P(♥ | ace) = \frac{P[♥ \cap ace]}{P[ace]} = \frac{P[A♥]}{P[A]} = \frac{\frac{1}{51}}{\frac{4}{51}} = \frac{1}{4}$, exactly as in the previous example.

EXAMPLE: In the situation of the previous example, find P(♠| 8).

SOLUTION: $P(♠ | 8) = \frac{P[♠ \cap 8]}{P[8]} = \frac{P[8♠]}{P[8]} = \frac{\frac{1}{51}}{\frac{3}{51}} = \frac{1}{3}$. For this problem, the missing card is relevant.

Here is another conditional probability example:

EXAMPLE: An urn has 7 red balls and 3 green balls. Suppose that you take two balls without replacement. What is the probability that both are red?

SOLUTION: The event we need can be described as $R_1 \cap R_2$. Do this as follows:

$$P(R_1 \cap R_2) = P(R_1) \times P(R_2 | R_1) = \frac{7}{10} \times \frac{6}{9} = \frac{7}{15} \approx 0.4667$$

Now a tricky example...

EXAMPLE: Suppose that you deal two cards from a standard deck. What is the probability that the cards are the same suit?

SOLUTION: Let SS denote the event "same suit." Now divide up the sample space according to the suit of the first card; call the events as $♠_1$, $♥_1$, $♦_1$, and $♣_1$. Then

$$P(SS) = P(SS \cap ♠_1) + P(SS \cap ♥_1) + P(SS \cap ♦_1) + P(SS \cap ♣_1)$$

$$= P(♠_1)\, P(SS | ♠_1) + P(♥_1)\, P(SS | ♥_1) + P(♦_1)\, P(SS | ♦_1) + P(♣_1)\, P(SS | ♣_1)$$

$$= \left\{\frac{1}{4} \times \frac{12}{51}\right\} + \left\{\frac{1}{4} \times \frac{12}{51}\right\} + \left\{\frac{1}{4} \times \frac{12}{51}\right\} + \left\{\frac{1}{4} \times \frac{12}{51}\right\} = \frac{12}{51} = \frac{4}{17} \approx 0.2353$$

Note that you interpret $P(SS | ♠_1)$ as the probability of getting a spade on the second draw, *given* that you got a spade on the first draw.

Some people get to the answer immediately by observing that, no matter which suit is selected on the first draw, the conditional probability of getting another of the same suit on the second draw must be $\frac{12}{51}$. This clever trick would not work for a deck with a missing card!

⋖⋖⋖⋖EXPECTED VALUE EXAMPLES⋗⋗⋗⋗

These examples illustrate the concept of expected value.

EXAMPLE: A grab bag contains 20 "prizes" in identical boxes. Of the "prizes,"

12 have a value of \$ 2

6 have a value of \$ 5

1 has a value of \$10

1 has a value of \$20

What is your mathematical expectation if you select one of the "prizes?" How would feel about paying \$8 to get to select one of these?

SOLUTION: The simplest thing to note is that the total prize value of $12 \times \$2 + 6 \times \$5 + \$10 + \$20 = \$84$ is spread out over 20 tickets, so the mathematical expectation must be $\frac{\$84}{20} = \$4.20$. It seems silly to pay \$8 to play this game; if the proceeds go to charity you might be willing.

You can also note that

the probability is 0.60 that you get \$ 2
the probability is 0.30 that you get \$ 5
the probability is 0.05 that you get \$10
the probability is 0.05 that you get \$20

The mathematical expectation is then found as

$$\{ 0.60 \times \$2 \} + \{ 0.30 \times \$5 \} + \{ 0.05 \times \$10 \} + \{ 0.05 \times \$20 \}$$

which comes to the same \$4.20.

EXAMPLE: The roulette wheel has 38 compartments, of which 18 are red, 18 are black, and two are green. As the wheel is spun, a small metal ball bounces around until it settles in one of the compartments. If you bet \$1 on red at roulette, you will get nothing back with probability $\frac{20}{38}$, as this happens when black or green comes up. You will get two dollars back with probability $\frac{18}{38}$ when red comes up. Find the expected amount you will get back.

SOLUTION: You get back \$0 with probability $\frac{20}{38}$, and you get back \$2 with probability $\frac{18}{38}$. The expected amount that you get back is

$$\left\{ \$0 \times \frac{20}{38} \right\} + \left\{ \$2 \times \frac{18}{38} \right\} = \$ \frac{36}{38} \approx \$0.947$$

This is of course less than the \$1 you paid to play the game.

EXAMPLE: If you bet \$1 on red at roulette, you will lose that dollar with probability $\frac{20}{38}$. Your dollar will be returned to you along with one more dollar if red comes up. Find your expected return.

SOLUTION: This is the same as the previous problem, except that we are incorporating the bet into the arithmetic directly. Your return will be -\$1 with probability $\frac{20}{38}$ and will be +\$1 with probability $\frac{18}{38}$. This leads to an expected return of

$$\left\{ (-\$1) \times \frac{20}{38} \right\} + \left\{ (+\$1) \times \frac{18}{38} \right\} = -\$ \frac{2}{38} \approx \text{-\$ } 0.053$$

The accounting is completely consistent, as this is the same as the difference between \$1 and \$0.947 of the previous problem.

EXAMPLE: If you bet \$1 on number “28” at roulette, you will lose that dollar with probability $\frac{37}{38}$. Your dollar will be returned to you along with 35 others if “28” comes up, and this will happen with probability $\frac{1}{38}$. Find your expected return.

SOLUTION: Your return will be -\$1 with probability $\frac{37}{38}$ and will be +\$35 with probability $\frac{18}{38}$. This leads to an expected return of

$$\left\{ (-\$1) \times \frac{37}{38} \right\} \quad + \quad \left\{ (+\$35) \times \frac{1}{38} \right\} \quad = \quad -\$\,\frac{2}{38} \;\approx\; \text{-\$ } 0.053$$

This is the expected return of every bet that you can make at roulette. Not only is this game efficient at separating you from your money (at a rate of 5.3¢ for every dollar bet), it is also very boring.

EXAMPLE: In a certain gambling game, three dice are rolled. You pay \$1 to play, and you bet on the number "4."

If "4" does not comes up, you lose your dollar.
If "4" comes up once, you get back your dollar with one other.
If "4" comes up twice, you get back your dollar with two others.
If "4" comes up three times, you get back your dollar with three others.

Find your expected return.

SOLUTION: Note that the returns for these lines are -1, +1, +2, and +3. The probabilities are, respectively, $\frac{125}{216}$, $\frac{75}{216}$, $\frac{15}{216}$, and $\frac{1}{216}$. These can be computed using the binomial probability law, but that's another story. The expected return is

$$\left\{ (-1) \times \frac{125}{216} \right\} + \left\{ 1 \times \frac{75}{216} \right\} + \left\{ 2 \times \frac{15}{216} \right\} + \left\{ 3 \times \frac{1}{216} \right\} \quad = \quad -\frac{17}{216} \;\approx \text{-0.079}$$

This represents an expected loss of 7.9¢ for each dollar bet. This is much worse than roulette.

EXAMPLE: The probability that a medical test will correctly detect the presence of a certain disease is 98%. The probability that this test will correctly detect the absence of the disease is 95%. The disease is fairly rare, found in only 0.5% of the population. If you have a positive test (meaning that the test says "yes, you got it") what is the probability that you really have the disease?

The best way to deal with such problems is to apply the proportions exactly to a large population. Say that you have 100,000 people. With the given facts, 0.5% of these, or $0.005 \times 100{,}000 = 500$ actually have the disease. Our state of information is this:

	Test POSITIVE	Test NEGATIVE	TOTAL
Have disease? YES			500
Have disease? NO			99,500
TOTAL			100,000

We're told that the test will correctly detect the presence in 98% of the people who actually have the disease. Thus, we expand our information to this (using $98\% \times 500 = 490$ and getting the 10 by subtraction):

	Test POSITIVE	Test NEGATIVE	TOTAL
Have disease? YES	490	10	500
Have disease? NO			99,500
TOTAL			100,000

We are also told that the test will correctly note the absence of the disease in 95% of the people who don't have the disease. Since $95\% \times 99{,}500 = 94{,}525$, we complete the table as follows:

	Test POSITIVE	Test NEGATIVE	TOTAL
Have disease? YES	490	10	500
Have disease? NO	4,975	94,525	99,500
TOTAL	5,465	94,535	100,000

Of course, in this final step we can note the column totals also.

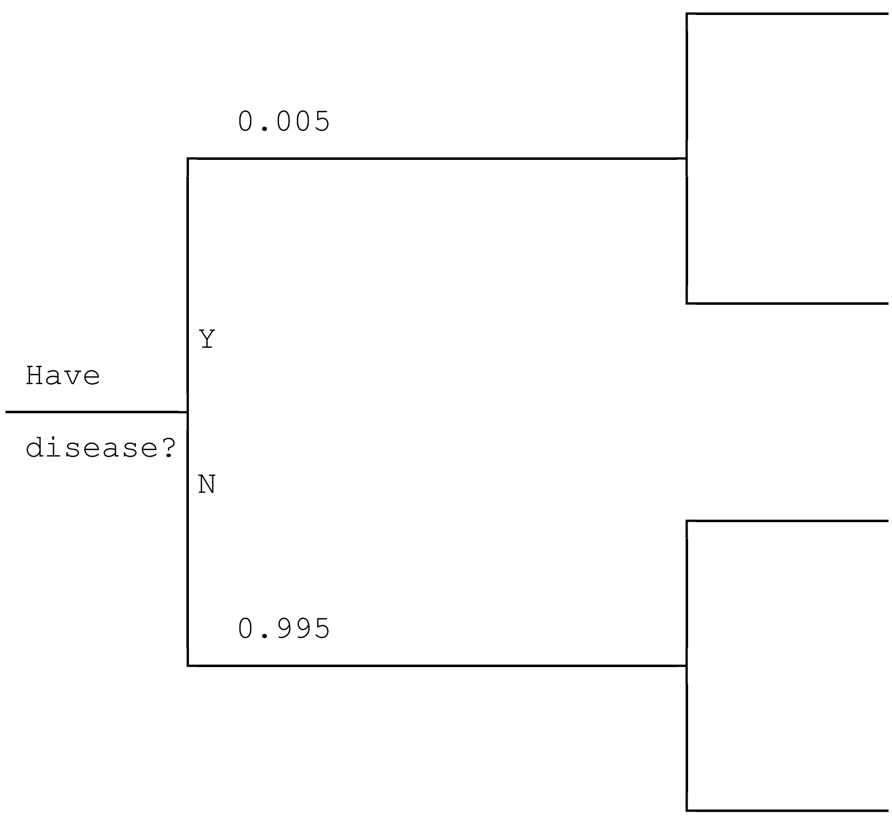
Conditional on a positive test, what's the probability of actually having the disease? We see that out of 5,465 persons showing positive on the test, only 490 have the disease. Our probability is then $\frac{490}{5,465} \approx 0.0897 \approx 9\%$. This is PV+ = P($D$ | T+) = predictive value positive. You should compare this to 0.5%, the probability of having the disease if the test is not done at all.

Conditional on a negative test, what's the probability of NOT having the disease? It's $\frac{94,525}{94,535} \approx 0.999894$. This is PV- = P($D'$ | T-) = predictive value negative. This should be compared to 0.995, the probability of not having the disease if the test is not done at all.)

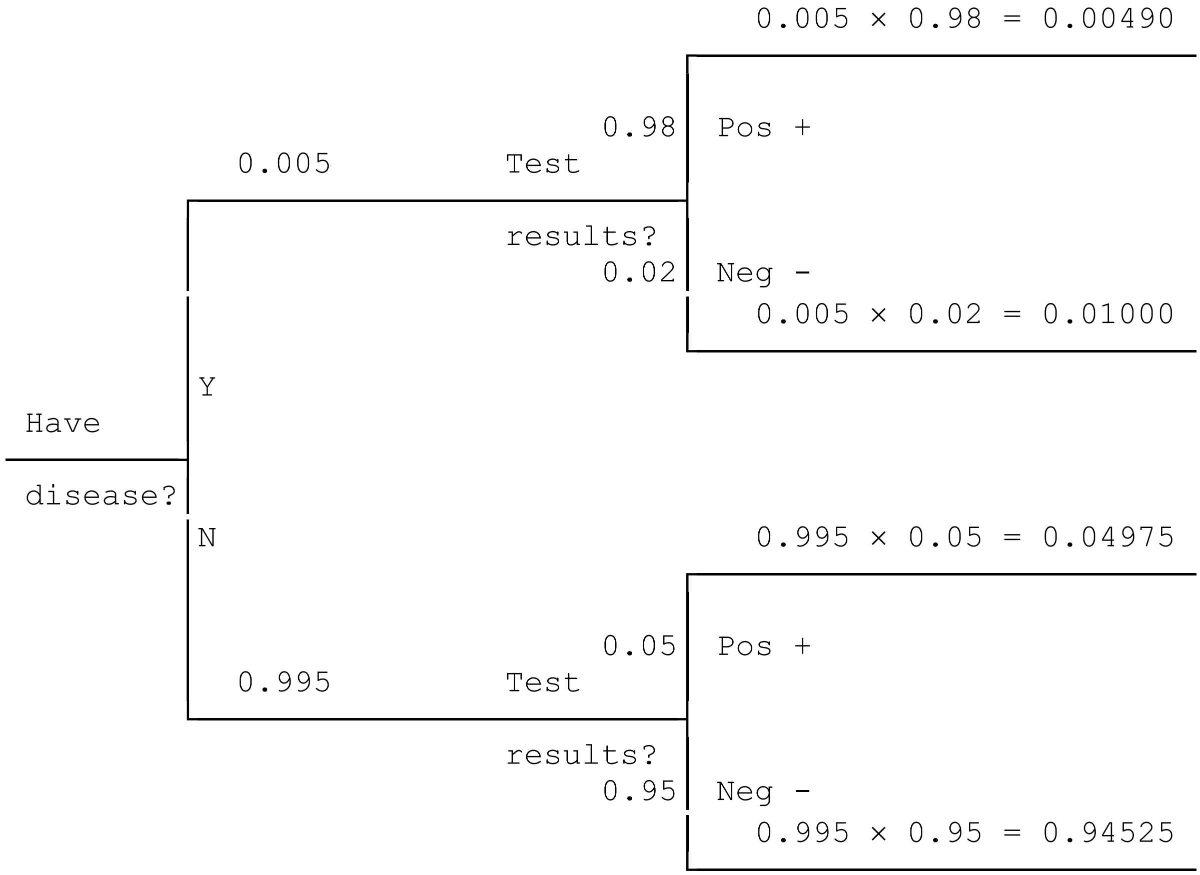
Also, P(T+ | D) is called sensitivity, while P(T- | D') is called specificity.

This is highly entertaining in the medical context, but it also applies directly to screening for uncommon industrial defects.

This medical screening example handout can be done through a probability tree. Start with the following:



At this point, you can extend with the probabilities for the tests.



Now, the probability of a positive test is 0.00490 + 0.04975 = 0.05465. Of these, the proportion who actually have the disease is 0.00490, and thus the conditional probability of having the disease, given a positive test, is $\frac{0.00490}{0.05465} \approx 0.0897 \approx 9\%$.

The arithmetic is identical, but the organization of the work is different.

There are some serious difficulties with the probability tree method. Here we made the first division in the tree based on “Have disease? (Y/N).” If we had made the first division on “Test results? (Pos +/Neg -)” then we would have had a lot of trouble!

<<<<<<<<<<<<<<<<<EFRON’S DICE>>>>>>>>>>>>>>>>

A gambler presents you with three dice, whose sides contain the numbers as indicated.

Die *A*:	3	9	10	14	28	40
Die *B*:	1	5	11	16	30	45
Die *C*:	2	8	13	19	20	31

A game is offered to you. You will select one of the three dice. The gambler will then select one of the other two. You each roll your chosen die. The person with the higher number showing will then collect $5 from the person with the lower number. Suppose that you decide to play this game. What is your expected return?

Suppose that die *A* is used against die *B*. You can show that $P(A > B) = \frac{17}{36}$,

$P(B > A) = \frac{19}{36}$, so die *B* is preferred over die *A*.

Suppose that die *B* is used against die *C*. You can show that $P(B > C) = \frac{17}{36}$,

$P(C > B) = \frac{19}{36}$, so die *C* is preferred over die *B*.

Now suppose that die *C* is used against die *A*. You can show that $P(C > A) = \frac{17}{36}$,

$P(A > C) = \frac{19}{36}$, so die *A* is preferred over die *C*.

Well. . . you know now how the gambler will play the game. If you select die *A*, he will select die *B*; if you select die *B*, he will select die *C*; if you select die *C*, he will select die *A*.

This shows that the ordering of the dice is *nontransitive*. This means that *A* can be beaten (on average) by *B*, *B* can be beaten by *C*, and *C* can be beaten by *A*. This phenomenon cannot happen for population means or expected values, which must obey a transitive ordering.

This example is due to Bradley Efron (though not with these numbers), and the game is sometimes called “Efron’s Dice.”

⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖⋖EFRON’S DICE⋗⋗⋗⋗⋗⋗⋗⋗⋗⋗⋗⋗⋗⋗⋗⋗

A variant on Efron’s Dice appeared in the January 1998 *Esquire* magazine, pages 104-105. In the *Esquire* version there are four dice, which we’ll call *W*, *X*, *Y*, and *Z*. (These are not the letters used in the magazine article.) Here are the numbers on those dice:

Die *W*: 1 1 1 5 5 5
Die *X*: 2 2 2 2 6 6
Die *Y*: 3 3 3 3 3 3
Die *Z*: 0 0 4 4 4 4

It is easy to show that

$P(X > W) = \frac{2}{3}$ so that *X* is better than *W*

$P(Y > X) = \frac{2}{3}$ so that *Y* is better than *X*

$P(Z > Y) = \frac{2}{3}$ so that *Z* is better than *Y*

$P(W > Z) = \frac{2}{3}$ so that *W* is better than *Z*

Is this a better example?

* It is very simple to work out the probabilities.
* The probabilities are rather far from $\frac{1}{2}$.

This example, however, uses four dice rather than three.

By the way, there are some unused comparisons. You might check that

$P(W > Y) = \frac{1}{2}$

$P(X > Z) = \frac{5}{9}$

<<<<<<<<<<<<<< LET'S MAKE A DEAL>>>>>>>>>>>>>>

This example, based on the actual television show *Let's Make a Deal*, has caused considerable controversy. The host of this television show was Monte Hall, and his name frequently gets dragged into the discussion.

The situation is this. You are a contestant on a television game show, and you've reached the phase of the game in which you choose one of three unopened boxes. You get the prize inside the chosen box, and you've been informed that one prize is valuable (keys to a new car) while the other two are petty (turnips).

You choose one of the boxes. Let's assume that you've chosen box number 1. The game show's host now points out that certainly at least one of the unchosen boxes contains a turnip. With an emotional voice and theatrically-polished gesturing, he opens box number 2 to reveal a turnip. It's now clear that the car keys are either in the box you've chosen (number 1) or in box number 3.

The host now offers you an interesting choice. You can keep box number 1 *or you can switch your choice to box number* 3. What do you do?

There are many, many ways to rationalize this situation. The clean solution requires an enumeration of the sample space, along with the probabilities. Let's make, however, the following assumptions:

> The probability that the car keys were placed originally in box number 1 is $\frac{1}{3}$. This probability also applies to boxes 2 and 3.
>
> The show's host knows which box contains the car keys.
>
> The show's script requires the host to reveal exactly one unchosen box. That is, the show's host cannot arbitrarily decide whether or not to reveal an unchosen box. The host is committed in advance to getting you to the point where you will be offered the choice to switch.
>
> If the contestant originally selects the box with the car keys, then the host must decide which of the other two boxes to open. Assume that this selection is made with probabilities $\frac{1}{2}$ each.
>
> The probability mechanisms are symmetric in the boxes. That is, the logic we use when you've chosen box number 1 will also apply to a contestant who chooses box number 2 or box number 3.

Based on this organization, we lay out the facts as follows. Remember that you've chosen box number 1.

Car keys in box	Host opens box	Prob.	Explanation
3	2	$\frac{1}{3}$	The probability is $\frac{1}{3}$ that the car keys are in box number 3. The host's actions are forced; he must open box number 2 to keep the game going.
2	3	$\frac{1}{3}$	The probability is $\frac{1}{3}$ that the car keys are in box number 2. The host's actions are forced; he must open box number 3 to keep the game going.
1	2	$\frac{1}{6}$	The probability is $\frac{1}{3}$ that the keys are in box number 1, so that the probabilities for these lines must sum to $\frac{1}{3}$.
1	3	$\frac{1}{6}$	We've assumed the host will divide his probabilities evenly between the two boxes with turnips.

The game show host, however, has exposed box number 2. We now proceed to do this as a conditional probability exercise.

$$P[\text{ car keys in box 1 } | \text{ host opens box 2 }]$$

$$= \frac{P\left[\text{car keys in box 1} \cap \text{host opens box 2}\right]}{P\left[\text{host opens box 2}\right]}$$

$$= \frac{P\left[\text{car keys in box 1}\right]P\left[\text{host opens box 2}\,\middle|\,\text{car keys in box 1}\right]}{P\left[\text{host opens box 2}\right]}$$

$$= \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{1}{2}} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

The numerator here is the probability on line 3 above, and the denominator is the sum of the probabilities on lines 1 and 3.

The tricky detail is that P[host opens box 2] = $\frac{1}{2}$. Our assumptions above indicate that he must open a box and, once we've selected box 1, the probabilities have to come down equally on box 2 and box 3.

It follows by subtraction that the probability that the keys are in box 3 is now $\frac{2}{3}$. We can explore this directly:

$$\text{P[car keys in box 3 | host opens box 2]}$$

$$= \frac{\text{P}\left[\text{car keys in box 3} \cap \text{host opens box 2}\right]}{\text{P}\left[\text{host opens box 2}\right]}$$

$$= \frac{\text{P}\left[\text{car keys in box 3}\right]\text{P}\left[\text{host opens box 2} \middle| \text{car keys in box 3}\right]}{\text{P}\left[\text{host opens box 2}\right]}$$

$$= \frac{\frac{1}{3} \times 1}{\frac{1}{2}} = \frac{2}{3}$$

Some people use the following points to give a simple argument in favor of switching boxes:

> If you keep the box you’ve chosen first, the probability is $\frac{1}{3}$ that you get the car keys. This is just the probability of an initial correct guess.
>
> If you switch away from the box you’ve chosen first, the probability is $\frac{1}{3}$ that you’ve given up the correct box. Thus the probability is $\frac{2}{3}$ that you have made a move that wins the car keys.

Why is this so controversial? On the first pass through the problem, you seem to feel that you were confused about the mathematics — but seeing the careful explanation makes you sure of the correctness of the switching strategy. After a while, however, you realize that the game is resting on unstated premises and vague assumptions by the game-player. Specifically, we’ve assumed that the host’s behavior in revealing a turnip-filled box is part of a prespecified plan (along with all the other things listed as assumptions above). What if these assumptions just aren’t correct?

Here is a similar-sounding problem, but it’s not quite the same. Suppose that you are taking an exam with a multiple-choice question that has three options. You have no idea as to which option is correct, so you simply decide to guess on (a). After the exam has been going for about 20 minutes, the proctor announces, “There is a typographical error on (c), so I have to tell you that (c) is incorrect.” Would you change your answer from (a) to (b)?

<<<<<<<<<<<<<<THE PROSECUTOR’S FALLACY >>>>>>>>>>>>>>

This document draws from the article “Interpretation of Evidence, and Sample Size Determination,” by C. G. G. Aitken. This appears in the collection *Statistical Science in the Courtroom*, edited by Joseph Gastwirth.

A defendant is standing trial for a crime. We’ll use the following symbols:

G denotes the event “the defendant is guilty of the crime.”
I denotes the event “the defendant is innocent of the crime.”
E denotes the evidence (which could include fingerprints, witness testimony, or DNA samples)

It is considered very important to determine $P(I \mid E)$, the probability that the defendant is innocent, given the evidence.

The value of $P(E \mid I)$, the probability that this evidence could come from an innocent defendant, is not directly relevant. Mixing $P(I \mid E)$ and $P(E \mid I)$ is called the prosecutor’s fallacy. This simple example illustrates this point.

In a town served by blue taxis and green taxis, there is a hit-and-run accident caused by a taxi. A single fallible witness sees the accident and reports that the taxi is green. As a result, the green taxi company becomes the defendant. Thus $P(E \mid I)$ represents the probability of the witness’s report *given* that the green company is innocent; this is not directly relevant, as it merely says something about the quality of the witness. The very relevant quantity is $P(I \mid E)$, which is the probability that the green company is innocent, given the evidence.

Scientific testimony will provide $P(E \mid I)$ and $P(E \mid G)$. That is, the scientist (or expert witness) can assess the probabilities that the evidence would arise from either an innocent person or a guilty person.

The critical missing information is $P(I)$, or equivalently $P(G)$. We need to know the probability that, with no information about evidence, the person is innocent. For example, if the defendant only rarely goes to the location in which the crime occurred, then $P(I)$ would be large.

The completed version of the calculation uses Bayes’ theorem:

$$P(G \mid E) = \frac{P(E \mid G)\,P(G)}{P(E)} = \frac{P(E \mid G)\,P(G)}{P(E \mid G)\,P(G) + P(E \mid I)\,P(I)}$$

A logically equivalent form is this:

$$P(I \mid E) = \frac{P(E \mid I)\,P(I)}{P(E)} = \frac{P(E \mid I)\,P(I)}{P(E \mid G)\,P(G) + P(E \mid I)\,P(I)}$$

If the first equation is divided by the second, we produce this:

$$\frac{\mathrm{P}(G \mid E)}{\mathrm{P}(I \mid E)} = \frac{\mathrm{P}(E \mid G)}{\mathrm{P}(E \mid I)} \times \frac{\mathrm{P}(G)}{\mathrm{P}(I)}$$

The left side represents the odds on guilt, given the evidence. This should determine the decision.

The first factor on the right can be determined by a scientist. This factor represents the relative likelihood of the evidence under the two scenarios.

The second factor on the right is simply the odds that the defendant is guilty, without considering the evidence. Equivalently, you can say that this odds is determined from all the other forms of information.

A related fallacy comes from the defense side. Suppose that there are M people who had the knowledge and the opportunity to commit a particular crime. This set of people might, for example, consist of all M employees of a company who knew that the daily cash was taken to the bank in a green Buick around 9:30 p.m. The defense lawyer will then insist that the probability that the defendant committed the crime is $\frac{1}{M}$. This is simplistic logic because it ignores all other information. Let $p_1, p_2, \ldots, p_M$ be the probabilities of guilt for these M people. That is, we have p_i = P(person i is guilty). We must have $\sum_{i=1}^{M} p_i = 1$ and with no other facts, we would seem to have $p_i = \frac{1}{M}$ for each person. However, there *is* other evidence E, and the task is to assess

P(person i is guilty | E)

for each i. After considering E, we would have

$$\sum_{i=1}^{n} \mathrm{P}(\text{person } i \text{ is guilty} \mid E) = 1$$

There is no requirement that each summand be exactly $\frac{1}{M}$.