probability that switching will win, conditional on Mr. Barker having chosen to place the prize behind a particular door. If the prize is behind door 1, switching wins in exactly the cases where we choose doors 2 or 3. This has probability \( \frac{2}{3} \) for each door, so
\[
P(\text{switching wins}) = P(1)P(\text{win} \mid 1) + P(2)P(\text{win} \mid 2) + P(3)P(\text{win} \mid 3) = P(1) \times \frac{2}{3} + P(2) \times \frac{2}{3} + P(3) \times \frac{2}{3} = \frac{2}{3}
\]
no matter how Mr. Barker chooses to place the prize, where \( P(1) \) is his probability of placing it behind door 1, etc. We are using probability as a logical engine without really exhibiting any model at all. If there is no model in existence embodying our assumptions (i.e. if our assumptions are mutually contradictory) we can, and likely will, get an answer that is false. If our assumptions fall short of being sufficient to determine a unique answer then we may flail around using the logical engine approach, although some interesting partial information may come of it.

9. **INDEPENDENT EVENTS.** We are familiar with the notion that smoking increases the risk of particular cancers. In probabilistic terms \( P(\text{cancer} \mid \text{smoker}) > P(\text{cancer}) \). The usual interpretation would be that \( P(\text{cancer}) \) is the probability that a person selected at random from a given population will develop the cancer. This is less than \( P(\text{cancer} \mid \text{smoker}) \) which is the chance that a person selected from those of the population who smoke will develop the cancer.

On the other hand, we hope and trust that the usual level of fluoride exposure from toothpastes and drinking water has no bearing on cancer. We expect that \( P(\text{cancer} \mid \text{fluoride}) = P(\text{cancer}) \), at least approximately. In such a case we say that the events "cancer" and "fluoride" are independent. Here are two equivalent definition of independence between events \( A, B \) with \( P(A) > 0 \)
\[
A, B \text{ independent if and only if } P(B \mid A) = P(B)
\]
\[
A, B \text{ independent if and only if } P(A \cap B) = P(A)P(B)
\]
If \( P(A) = 0 \) it is independent of any other event by definition. This makes sense since any event that cannot occur tells us nothing about another event. This agrees with the second formulation but the first cannot be defined for \( P(A) = 0 \). Use the definition of conditional probability to prove that these two definitions are equivalent as are the following
\[
A, B \text{ independent}
\]
\[
\overline{A}, \overline{B} \text{ independent}
\]
and so forth. What we are saying is that whether or not one of independent events occurs has no bearing on the probability of the other or its complement.

**VENN DIAGRAM FOR INDEPENDENT EVENTS.** Suppose for example that we know \( P(\text{rain tomorrow in East Lansing}) = 0.4 \) and \( P(\text{Yuan catches a fish in Shanghai tomorrow}) = 0.05 \). It seems far fetched that we would be led to revise our probability for Yuan based upon whether or not it rains in East Lansing tomorrow. Perhaps it is safe to regard these events as being independent. Since generally
\[
P(\text{rain and catch}) = P(\text{rain})P(\text{catch} \mid \text{rain})
\]
if these events are independent we have \( P(\text{catch} \mid \text{rain}) = P(\text{catch}) \) so the above becomes
\[
P(\text{rain and catch}) = P(\text{rain})P(\text{catch}) = 0.4 \times 0.03 = 0.012.
\]
It then follows also that
\[
P(\text{rain or catch}) = P(\text{rain}) + P(\text{catch}) - P(\text{rain and catch}) = 0.4 + 0.03 - 0.012 = 0.418.
\]
Complete the Venn diagram with this information. Note that you completing a Venn diagram based on three pieces of information: \( P(A), P(B) \), and the fact that \( A, B \) are independent.
CONTINGENCY TABLE FOR INDEPENDENT EVENTS. Independence of events is seen in a probability contingency table as proportionality. Fill out the table for the given information. Note that the rows (columns) are proportional to one another. Proportionality is the appearance of independence in a contingency table.

<table>
<thead>
<tr>
<th></th>
<th>no catch</th>
<th>marginal totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain</td>
<td></td>
<td>0.4</td>
</tr>
<tr>
<td>no rain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>marginal totals</td>
<td></td>
<td>grand total = 1</td>
</tr>
</tbody>
</table>

10. Draws with replacement are independent. As before, consider draws from the urn

\[5 B, 10 W, 7 G\]

but this time with replacement (and equal probability). It is like dealing from a deck to which each card is returned, and the deck shuffled, before each new draw. Evaluate

\[P(B_1, B_2, G_3) = P(B_1)P(B_2 | B_1)P(G_3 | B_1, B_2)\]

\[P(B_2) = P(B_1)P(B_2 | B_1) + P(W_1)P(B_2 | W_1) + P(G_1)P(B_2 | G_1)\]

11. Risk accumulation. It seems reasonable that random failures, at least with proper use, are independent. Suppose the probability of failure with each use is 0.005 (hypothetical). What is the chance that there are no failures with 500 uses?

\[P(\text{no failure with 500 uses})\]

From (21) above describe why the e-approximation of this probability is

\[e^{-0.005 \cdot 500} = 0.082085 \text{ (rounded, check it)}\]

Is the e-approximation accurate in this case?
Does such a failure risk surprise you? On the web (e.g., site “Alice” of the Columbia University Health Center) you will see statements like the following (not an actual quote): 14% annual pregnancy rate for women with “regular” condom use; 9% annual pregnancy rate for women with “perfect” condom use. It’s hard to know what such statements mean because of all the variables involved but the order of magnitude is, one would think, high enough to draw some attention.

1.2. DISJOINT VS INDEPENDENT. The events A = “the last shuttle flight is in January” and B = “the last shuttle flight is in June” are disjoint (cannot both occur). Assuming that each is neither certain nor impossible are A, B independent?

What if P(A) = 0? Are they independent then?

1.9. FAMOUS PROBLEM THAT HELPED START PROBABILITY. Independence is at the root of the solution of a famous problem posed by Chevalier de Mere, to the mathematician Blaise Pascal. At issue was which was the better bet (a) getting at least one ace \( \bullet \) in four rolls of a single die or (b) getting at least one double ace \( \bullet \bullet \) (called snake eyes) in 24 rolls of two dice? It seemed from experience that (a) was slightly better but nobody had been able to prove it. Find out for yourself as follows:

In the roll of one die \( P(\bullet) = \frac{1}{6} \) so \( P(\) no aces in 6 rolls\) = \( \left( \frac{5}{6} \right)^6 \)

In the roll of two dice \( P(\bullet \bullet) = \frac{1}{36} \) so \( P(\) no snake eyes in 24 rolls\) = \( \left( \frac{35}{36} \right)^{24} \)

Calculate and justify the above. De Mere observed that \( 4 \times \frac{1}{6} \) is the same as \( 24 \times \frac{1}{36} \) but that has no bearing on the solution.

Which of (a) or (b) is best?

Note that \( \left( \frac{35}{36} \right)^{24} \) is approximately \( e^{-2} \) = 0.5134 so P(no snake eyes in 24 rolls) is approximately \( 1 - 0.5134 = 0.4865 \)...

1.4. PARALLEL AND SERIES RISK. The risk of failures in elaborate systems such as airframes or networks can be assessed by creating a layout of the logical interconnections of their constituent components. This is called risk analysis. Ultimately, the fundamental building blocks of such a layout are parallel (“or” junctures) or series (“and” junctures).

A good way to think about this is to imagine a system of two components in series.

\[ \begin{array}{c}
\text{A} \\
\text{B} \\
\text{SERIES}
\end{array} \]

If P(A does not fail) = 0.9, P(B does not fail) = 0.95, and failure of these components are independent events

P(series system does not fail) = P(A does not fail and B does not fail) = 0.9 \times 0.95 = 0.855.

Another use of the same components is to place them in parallel (think of two bridges over the same river).

\[ \begin{array}{c}
\text{A} \\
\text{B} \\
\text{PARALLEL}
\end{array} \]
P(parallel system fails) = P(A fails and B fails) = (1 - 0.9)(1 - 0.95) = 0.1 \times 0.05 = 0.005.

P(parallel system does not fail) = 1 - P(parallel system fails) = 1 - 0.005 = 0.995.

You can see that the series arrangement increases the risk of failure since all components in series need to work if the system is to work. On the other hand the parallel system reduces the risk of system failure since the system will work if even one component works.

A more complex system can be dealt with by "working inside out." You would first do the series A1 B1 then the series A2 B2 B3, then place those two series in parallel.

\[
\begin{array}{cccccc}
\vline & \vline & \vline & \vline & \vline & \vline \\
A1 & \_ & \_ & \_ & B1 & \_ \\
\_ & \_ & \_ & \_ & \_ & \_ \\
A2 & \_ & \_ & B2 & \_ & \_ \\
\_ & \_ & \_ & \_ & \_ & \_ \\
B3 & \_ & \_ & \_ & \_ & \_ \\
\_ & \_ & \_ & \_ & \_ & \_ \\
\end{array}
\]

Assume the components A1, A2 are like A above. Also assume the components B1, B2, B3 are like B above. Moreover assume all five of these components fail independently.

P(above 5 component system does not fail) =

These problems are vastly simplified if failures of these events are independent although general probability rules can be used if enough data is available on failures of dependent components.

15. THE LIGHT BULB PARADOX. Why do some light bulbs seem to last a very long time before burning out? This another of life's little mysteries that can be partially explained by probability. Some years ago I saw such a bulb in an old firehouse in Haslett, MI that had been burning since the firehouse was built. An even older one is reputed to exist in New York and there are doubtless others. Imagine a string of tiny heads each one of which independently "blow off" with small probability p.

For some bulbs, failure occurs when a power fluctuation occurs. Think of the heads as representing opportunities for a power fluctuation (blow off) to occur. If there has been no blow off by head n this does not affect the odds for future heads, so it is as if the bulb is "born again." That is, this model predicts that the lifetimes of all similar bulbs follows a distribution that is the same as the lifetimes of bulbs that have been in service for one year (or any other period of time). It is as though a bulb, having been found burning, is "born again," and has all of the expectations for future life that a new bulb does.

The probability that the first blow off occurs after head n is \((1 - p)^n = e^{-np}\) for \(n = \infty\) and \(p = 0\) suggesting an exponential decay for the plot of P1 bulb life \(t\) vs \(t > 0\). A seeming paradox is that, for any fixed time \(t_0\), the conditional probability

\[P(\text{bulb lives longer than } t_0 \text{ additional units of time } | \text{ bulb is still burning at time } t_0)\]

does not depend upon \(t_0\). So truly, under this model of failure by rare catastrophe, old bulbs act exactly as new.

Calculate P1 you wait more than 2 tosses for the first head.

Toss a coin until the first head occurs. Repeat 20 times. With what relative frequency do you find that the first head occurs after toss two? Compare with your calculation above.

INSPECTION PARADOX. The average gap between heads that blow off will on average be one half of the average gap between the nearest blow off points on either side of a fixed position on the string.
gap to left avg same as $\Theta \leftrightarrow \Theta$ does $\uparrow$ gap to right avg same as $\Theta \leftrightarrow \Theta$ does

fixed point

So if I am standing at the side of a road waiting to cross it will typically seem that an unusually large gap in traffic is at hand. This is of illusory advantage of course since the downstream part of the gap does me no good. Something similar may be at work when a driver looks to the other lane from time to time and sees generally large gaps. Another example concerns the boss who steps into the store occasionally and, enquiring how long it has been since the last customer, adds that to the wait for the next customer thinking what a long total time it is. Under this model it is because two gaps, past and future, each having the character of a gap between consecutive customers, have been added.

Toss a die until you get a second ace $\Box$. Note gap2aces = (1 + the number of tosses between the two aces). Repeat 20 times and average the results.

Toss a die 30 times. There will probably be an ace $\Box$ on either side of a mark placed between tosses 15 and 16. Note fixedgap = (1 + the number of tosses between these two aces), or the gap to toss 1 or 30 if there is no ace to one side or the other. Repeat 20 times and average these gaps. Do you get something on the order of twice what you get above?

16. MOST FAVORABLE CASINO GAME (CRAPS). Actually, there are counting schemes for Blackjack that enjoy a tiny advantage over the house, albeit too small to recommend their use unless elaborate teams of players cooperate in certain ways without being detected. So technically Craps (dice) is not the most favorable game. You can however play a Craps game which is nearly fair. If you are in a mood to win (or lose) big you can always wager one large stake at nearly even odds.

Ignoring all that, we admit to the fact that you pay for playing, by losing at predictable rates. So we're talking about finding a casino game where you can lose slowly, thus prolonging your losing fun and giving the law of averages a surer hand in seeing that you do lose. Craps offers these kinds of bets. So if you are roped into visiting a casino and are content to look like you know what you are doing, lose little or maybe even win a little, here is your game.

CRAPS. One person, the "shooter," rolls a pair of dice until their turn is finished, whereupon the dice are passed to another shooter. The shooter's first roll is called the "come out" roll.

PASS LINE BET. Anyone at the table may make the pass line bet by simply placing chips on the pass line. You will either lose that money or double it (double or nothing). You win outright if the shooter rolls a total of 7 or 11 on their come out roll. You lose outright if they roll 2 or 3 or 12. If they roll anything else (4, 5, 6, 8, 9, or 10) that becomes the "point." The shooter continues to roll until either their "point" comes up again or a 7 (craps) is thrown, whichever is first. If they "make their point" before throwing a 7 the pass line bet is won. Otherwise the pass line bet is lost. Let's calculate the probability that pass line bet is won.

$$P(\text{pass line is won}) = \sum_{k=2}^{12} P(\text{come out} = k) P(\text{pass line is won} | \text{come out} = k)$$

$$= P(2) P(\text{win} | 2) + P(3) P(\text{win} | 3) + \ldots + P(12) P(\text{win} | 12)$$

We have earlier calculated the distribution for the total of two dice

$$
\begin{array}{cccccccccccc}
  k & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
  P(k) & \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \\
\end{array}
$$

Some values of $P(\text{win} | k)$ are easy

$$
\begin{array}{cccccccccccc}
  k & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
  P(k) & \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \\
\end{array}
$$
The others, like $P_{\text{win}}(\mid 4)$ are harder. In this case it is assumed that the come out roll is a 4 and for you to win the shooter has to (again) roll 4 before 7. Conditionally, we are just rolling dice to get a 4 before 7:

$$P_{\text{win}}(\mid 4) = P(4 \mid 7).$$

To calculate it we employ a shortcut. Surely there will come a roll which decides the matter. It is the first roll when either 4 or 7 results. At that point the dice "don't know about the previous rolls" so to speak. In other words, if someone announces that the game is over, i.e. a 4 or 7 has been rolled, we don't care to ask how many rolls it took. We simply know that roll was 4 or 7 so

$$P(4 \mid 7) = \begin{pmatrix} P(4) \end{pmatrix} = \frac{\frac{3}{3} + \frac{3}{6}}{\frac{3}{3} + \frac{6}{6}} = \frac{3}{3}$$

$$P(5 \mid 7) = \begin{pmatrix} P(5) \end{pmatrix} = \frac{\frac{3}{5} + \frac{3}{7}}{\frac{3}{5} + \frac{7}{7}} = \frac{2}{5}$$

and so on.

For $P(4)$ we can use the law of total probability

$$P(4) = \begin{pmatrix} P(4) \end{pmatrix} = 1$$

$$P(5) = \begin{pmatrix} P(5) \end{pmatrix} = 3$$

Similarly $P(5)$:

$$P(5) = \begin{pmatrix} P(5) \end{pmatrix} = 3$$

$$P(6) = \begin{pmatrix} P(6) \end{pmatrix} = 3$$

$$P(1) = \begin{pmatrix} P(1) \end{pmatrix} = 1$$

The pass line bet, which pays even money, wins with probability 0.4929299... only slightly below 0.5.

DON'T PASS LINE BET. This bet is against the pass line bet except a come out roll of 12 is a "standoff." It is slightly more favorable than the pass line and if there is not an outright win or loss the don't pass line better may (in many Casinos) triple the amount of the wager, the additional double bet being paid at conditionally even odds. The additional chips are simply placed adjacent to the original bet. For example, if the coming out roll is 4 there is now a conditional 1/3 chance of a win. If the don't pass line better originally bet +$10 the additional +$20 bet will pay back +$30 if 4 is rolled before 7. The original +$10 don't pass line bet still stands and is unaffected by this additional bet. Less than +$20 may also be bet, but why hold anything at even odds in a casino?

The conditional probability of a win in games where 12 is not the come out roll is

$$P(\text{don't pass line bet wins} \mid \text{come out roll is not 12}) = \frac{P(\text{come out roll is not 12 and don't pass line wins})}{P(\text{come out roll is not 12})} - P(\text{come out roll is not 12 and pass line bet wins})$$

$$= \frac{P(\text{come out roll is not 12}) - P(\text{pass line bet wins})}{P(\text{come out roll is not 12})} = \frac{1}{2} + \frac{1}{9} = 0.4929299...$$

That is, the pass line bet by 0.0000577201 as well as allowing standoff on 12 and late bets at even odds!

17. BAYES' FORMULA. These days, probability is being used as a logical engine with which to process information.

probability is being used as a logical engine

This is because probabilities can be conveniently revised to reflect changing information. The tools used for this purpose are primarily the Law of Total Probability and a formula worked out by Englishman Rev. Thomas Bayes in 1761.
Consider an event \( B \) and partition \( A_1, A_2, \ldots, A_n \) of the sample space into disjoint events. Think of a situation in which we already have a-priori knowledge of probabilities
\[
P(A_1), P(A_2), \ldots, P(A_n) \quad \text{(all known)}
\]
\[
P(B \mid A_1), P(B \mid A_2), \ldots, P(B \mid A_n) \quad \text{(all known)}
\]
i.e. we know the probabilities for \( n \) different possible "causes" \( A_k \) and we also know the probabilities with which each of these possible "causes" leads to an occurrence of \( B \). If event \( B \) is then seen to occur we can revise our probabilities according to Bayes' Formula
\[
P(A_k \mid B) = \frac{P(A_k) P(B \mid A_k)}{P(A_1) P(B \mid A_1) + P(A_2) P(B \mid A_2) + \ldots + P(A_n) P(B \mid A_n)}
\]

**OIL EXAMPLE.** In this example we have a potential oil field. An initial survey suggests
\[
P(\text{OIL}) = 0.7 \quad P(\text{NO OIL}) = 0.3
\]
i.e. \( A_1 \) \( A_2 \)

These are our "causes." We now perform a test which is designed to detect oil. The test comes back "+" or "−" indicating oil or no oil. The test is not perfect however, so false positives and false negatives may occur. Our consultants estimate the following performance probabilities for the output of this test
\[
P(+) \mid \text{OIL} = 0.9 \quad P(+) \mid \text{NO OIL} = 0.2
\]
i.e. \( B \) \( A_1 \) \( B \) \( A_2 \)

Notice that the probabilities 0.9 and 0.2 do not add to one. They are not probabilities of different contingencies that exhaust the sample space. Rather they are probabilities of a fixed event "+" under different contingencies. Bayes' Formula applied to this information gives
\[
P(\text{OIL} \mid +) = \frac{P(\text{OIL}) P(+) \mid \text{OIL}}{P(+) \mid \text{OIL} P(\text{OIL}) + P(+) \mid \text{NO OIL} P(+) \mid \text{NO OIL}}
\]
\[
= \frac{0.7 \times 0.9}{0.7 \times 0.9 + 0.3 \times 0.2} = \frac{0.63}{0.69} = 0.913
\]

This means that, if the test comes back positive, the probability of OIL is increased from its a-priori value 0.7 to its a-posteriori value \( P(\text{OIL} \mid +) = 0.913 \).

What do you think will happen to the probability of OIL if a negative test result is returned? Work it out using Bayes' formula
\[
P(\text{OIL} \mid -)
\]

**PROOF OF BAYES’ FORMULA.** By multiplication of probabilities
\[
P(A_k) P(B \mid A_k) = P(A_k \cap B)
\]
and by total probability
\[
P(A_1) P(B \mid A_1) + P(A_2) P(B \mid A_2) + \ldots + P(A_n) P(B \mid A_n)
\]
\[
= P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_n \cap B) = P(B)
\]
so the quotient appearing in Bayes' Formula is
\[
\frac{P(A_k \cap B)}{P(B)} = P(A_k \mid B)
\]
18. **TREE DIAGRAM.** Bayes' calculations are sometimes facilitated by a tree diagram which can also help clarify the logic involved.

Above the branch "OIL" write \( P(\text{OIL}) = 0.7 \). Above "OIL +" write \( P(+ \mid \text{OIL}) = 0.9 \). Likewise below "OIL —" write \( P(— \mid \text{OIL}) = 0.1 \).

Notice that \( P(+) \mid \text{OIL}) + P(— \mid \text{OIL}) = 1 \). Continue filling out the tree, writing \( P(\text{NO OIL}) = 0.3 \) below "NO OIL," and so forth for \( \text{NO OIL}^+, \text{NO OIL} — \).

The multiplication rule tells us that \( P(\text{NO OIL} \text{ and } +) = P(\text{NO OIL} \mid +) \times P(+ \mid \text{NO OIL}) = 0.3 \times 0.2 = 0.06 \). Place this 0.06 at the tip of the branch from "NO OIL" through "+".

Likewise fill out the other three branch tips. The total of the branch tips is the total probability of the four disjoint regions of the Venn diagram and must equal one. If it does not total one check your work.

**BAYES FROM TREE.** If you have the tree diagram then Bayes’ probability \( P(\text{OIL} \mid +) \) is easily obtained as follows

\[
P(\text{OIL} \mid +) = \frac{\text{total of all tree tip probabilities through OIL and +}}{\text{total of all tree tip probabilities through +}}
\]

This general principle applies to all conditional or unconditional probability calculations no matter what the number of branches around the various nodes or the number of nodes of a tree diagram. After all, you are just adding up all of the disjoint pieces of a Venn diagram compatible with the numerator and denominator of whatever conditional probability you are calculating (unconditional probability is just conditional probability given \( SI \)).

19. **FALSE POSITIVE PARADOX.** Suppose \( P(\text{diseased}) = 0.01 \) (rare) and

\[
P(+) \mid \text{diseased} = 1 \quad \text{(the test certainly positive if person is diseased)}
\]

\[
P(+) \mid \text{not diseased} = 0.01 \quad \text{(the test infrequently positive if not diseased)}
\]

Looks like a reliable test! Combine this information (i.e., use probability as a "logical engine") to calculate \( P(\text{diseased} \mid +) \) the rate at which persons who test positive are diseased.

Are you surprised by this result? It comes about because

\[
P(\text{diseased} \mid +) = \frac{P(\text{diseased}) \times P(+) \mid \text{diseased}}{P(\text{diseased}) \times P(+) \mid \text{diseased} + P(\text{not diseased}) \times P(+) \mid \text{not diseased}}
\]

is small owing to the large value of \( P(+) \mid \text{not diseased} = 0.99 \), even though \( P(+) \mid \text{diseased} = 1 \) and \( P(+) \mid \text{not diseased} = 0.01 \) are fine. If this were hospital data it could happen if there is a public health scare that causes lots of healthy people to come in for a test "just to be sure." There is nothing wrong with the test per se, it is just being overwhelmed. If the hospital carefully screens patients they may be able to set \( P(\text{diseased}) \) down to something like 0.2. Then

\[
P(\text{diseased} \mid +) = \frac{P(\text{diseased}) \times P(+) \mid \text{diseased}}{P(\text{diseased}) \times P(+) \mid \text{diseased} + P(\text{not diseased}) \times P(+) \mid \text{not diseased}}
\]

\[
P(\text{diseased} \mid +) = 0.2 \times 1 \quad 0.2 \times 0.99 \times 0.01 = 0.716...
\]

\[
P(\text{diseased} \mid —) = \frac{P(\text{diseased}) \times P(—) \mid \text{diseased}}{P(\text{diseased}) \times P(—) \mid \text{diseased} + P(\text{not diseased}) \times P(—) \mid \text{not diseased}}
\]

\[
P(\text{diseased} \mid —) = \frac{0.2 \times 0.99 \times 0.01}{0.2 \times 1 \quad 0.2 \times 0.99 \times 0.01} = 0.013...
\]
SIMPSON'S PARADOX. This is closely related to Simpson's Paradox, a fine example being found in STATISTICS by Freedman, Pisani and Purves. It relates the fact that in a particular year the rate of admissions to graduate studies at Berkeley was better for male applicants than it was for female applicants (i.e. Pladmitted | male > Pladmitted | female). Candidates apply to programs, not to graduate study in general. When the data were broken down by program it was found that for every program k, except one where it was close to even, Pladmitted | male, program k > Pladmitted | female, program k. Complete reversal (of the overall finding) for every k, which is mathematically possible, was nearly achieved in this case. Apparently, women applied in relatively greater numbers than did men to programs that were the hardest for either sex to gain admission to. Through self-selection, women overwhelmed was appears to have been a fair admissions policy. In order to avoid such pitfalls, marketing studies, clinical trials and engineering studies, and so on, must block self-selection through such means as random assignment of subjects and other 'statistical experimental design' techniques.

20. A LOTTERY QUESTION. YOU DO NOT HAVE TO GAMBLE OR REGISTER FOR THE LOTTERY SITE TO COMPLETE THIS OPTIONAL ASSIGNMENT. Lottery prize structures remain something of a mystery, especially when they seem to go against common sense. One example is a New Jersey based lottery run on behalf of a consortium of states located in various regions of the U.S. The lottery is described on www.state.nj.us/lottery/games/1-1_big_game.shtml. Here is a portion of the basis description found there.

"MEGA MILLIONS...To play a board, you must select five numbers from 1 to 52 in the upper shaded box and select one (1) Gold Mega Ball number from 1 to 52 in the lower white box..."

For the “52 pick 5” players choose 5 different integers from 1 to 52. The prize for each player matching the winning five numbers is $175,000? Players choose any additional number 1 to 52 (Mega Ball) for a chance at sharing the Jackpot (now $12 million U.S.).

### Odds & Prizes

<table>
<thead>
<tr>
<th>Match</th>
<th>Prize*</th>
<th>Odds (per $1 play)</th>
</tr>
</thead>
<tbody>
<tr>
<td>🎥ólogo</td>
<td>$1,351,459,200</td>
<td>1 : 135,145,920</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$175,000</td>
<td>1 : 2,649,920</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$5,000</td>
<td>1 : 575,089</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$150</td>
<td>1 : 11,276</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$150</td>
<td>1 : 12,502</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$10</td>
<td>1 : 833</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$7</td>
<td>1 : 245</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$3</td>
<td>1 : 152</td>
</tr>
<tr>
<td>🎥ólogo +</td>
<td>$2</td>
<td>1 : 88</td>
</tr>
</tbody>
</table>

*Odds of winning jackpot are 1 : 135,145,920. Overall odds of winning are approximately 1 : 43.*  
*Subject to published rules of MEGA MILLIONS and the New Jersey Lottery, the fixed prize amounts indicated here may be pari-mutuel.*  
**The JACKPOT prize will be divided equally among multiple winners. The prize is paid in 26 annual installments unless Cash Option is selected."

As you see from the table, any winner of the 52 pick 5 can win a share of the Jackpot if they match the power ball too. The jackpot can run into six figures (mega millions, today $12 million U.S.) since it rolls over to the next lottery if no one claims it. Under the assumption that the lottery picks a set of 5 from 52 at random and independently chooses an integer from 1 to 52.
P(you win the 52 pick 5)

P(you win a share of the jackpot)

A QUESTION. Do you find the basic description clear about the payout structure? Do you think players pick numbers at random or are they more likely to pick particular numbers? What happens if everyone plays the same 52 pick 5 and they all win?

ANOTHER QUESTION. Does the chance of winning the jackpot, versus just winning the 52 pick 5, seem in line with the ratio between the big prize and $175,000? Why do you suppose the lottery sets it up with the given ratios?