

6.3 Conditional Probability and Independence

Conditional Probability

Two cubical dice each have a triangle painted on one side, a circle painted on two sides and a square painted on three sides. Applying the principal of inclusion and exclusion, we can compute that the probability that we see a circle on at least one top when we roll them is $1/3 + 1/3 - 1/9 = 5/9$. We are experimenting to see if reality agrees with our computation. We throw the dice onto the floor and they bounce a few times before landing in the next room.

Exercise 6.3-1 Our friend in the next room tells us both top sides are the same. Now what is the probability that our friend sees a circle on at least one top?

Intuitively, it may seem as if the chance of getting circles ought to be four times the chance of getting triangles, and the chance of getting squares ought to be nine times as much as the chance of getting triangles. We could turn this into the algebraic statements that $P(\text{circles}) = 4P(\text{triangles})$ and $P(\text{squares}) = 9P(\text{triangles})$. These two equations and the fact that the probabilities sum to 1 would give us enough equations to conclude that the probability that our friend saw two circles is now $2/7$. But does this analysis make sense? To convince ourselves, let us start with a sample space for the original experiment and see what natural assumptions about probability we can make to determine the new probabilities. In the process, we will be able to replace intuitive calculations with a formula we can use in similar situations. This is a good thing, because we have already seen situations where our intuitive idea of probability might not always agree with what the rules of probability give us.

Let us take as our sample space for this experiment the ordered pairs shown in Table 6.2 along with their probabilities.

Table 6.2: Rolling two unusual dice

TT	TC	TS	CT	CC	CS	ST	SC	SS
$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{4}$

We know that the event $\{\text{TT}, \text{CC}, \text{SS}\}$ happened. Thus we would say while it used to have probability

$$\frac{1}{36} + \frac{1}{9} + \frac{1}{4} = \frac{14}{36} = \frac{7}{18} \tag{6.13}$$

this event now has probability 1. Given that, what probability would we now assign to the event of seeing a circle? Notice that the event of seeing a circle now has become the event CC . Should we expect CC to become more or less likely in comparison than TT or SS just because we know now that one of these three outcomes has occurred? Nothing has happened to make us expect that, so whatever new probabilities we assign to these two events, they should have the same ratios as the old probabilities.

Multiplying all three old probabilities by $\frac{18}{7}$ to get our new probabilities will preserve the ratios and make the three new probabilities add to 1. (Is there any other way to get the three new probabilities to add to one and make the new ratios the same as the old ones?) This gives

us that the probability of two circles is $\frac{1}{9} \cdot \frac{18}{7} = \frac{2}{7}$. Notice that nothing we have learned about probability so far told us what to do; we just made a decision based on common sense. When faced with similar situations in the future, it would make sense to use our common sense in the same way. However, do we really need to go through the process of constructing a new sample space and reasoning about its probabilities again? Fortunately, our entire reasoning process can be captured in a formula. We wanted the probability of an event E given that the event F happened. We figured out what the event $E \cap F$ was, and then multiplied its probability by $1/P(F)$. We summarize this process in a definition.

We define the *conditional probability* of E given F , denoted by $P(E|F)$ and read as “the probability of E given F ” by

$$P(E|F) = \frac{P(E \cap F)}{P(F)}. \quad (6.14)$$

Then whenever we want the probability of E knowing that F has happened, we compute $P(E|F)$. (If $P(F) = 0$, then we cannot divide by $P(F)$, but F gives us no new information about our situation. For example if the student in the next room says “A pentagon is on top,” we have no information except that the student isn’t looking at the dice we rolled! Thus we have no reason to change our sample space or the probability weights of its elements, so we define $P(E|F) = P(E)$ when $P(F) = 0$.)

Notice that we did not prove that the probability of E given F is what we said it is; we simply defined it in this way. That is because in the process of making the derivation we made an additional assumption that the relative probabilities of the outcomes in the event F don’t change when F happens. This assumption led us to Equation 6.14. Then we chose that equation as our definition of the new concept of the conditional probability of E given F .⁴

In the example above, we can let E be the event that there is more than one circle and F be the event that both dice are the same. Then $E \cap F$ is the event that both dice are circles, and $P(E \cap F)$ is, from the table above, $\frac{1}{9}$. $P(F)$ is, from Equation 6.13, $\frac{7}{18}$. Dividing, we get the probability of $P(E|F)$, which is $\frac{1/9}{7/18} = \frac{2}{7}$.

Exercise 6.3-2 When we roll two ordinary dice, what is the probability that the sum of the tops comes out even, given that the sum is greater than or equal to 10? Use the definition of conditional probability in solving the problem.

Exercise 6.3-3 We say E is *independent* of F if $P(E|F) = P(E)$. Show that when we roll two dice, one red and one green, the event “The total number of dots on top is odd” is independent of the event “The red die has an odd number of dots on top.”

Exercise 6.3-4 Sometimes information about conditional probabilities is given to us indirectly in the statement of a problem, and we have to derive information about other probabilities or conditional probabilities. Here is such an example. If a student knows 80% of the material in a course, what do you expect her grade to be on a (well-balanced) 100 question short-answer test about the course? What is the probability that she answers a question correctly on a 100 question true-false test if she guesses at each question she does not know the answer to? (We assume that she knows what

⁴For those who like to think in terms of axioms of probability, we could give an axiomatic definition of conditional probability, and one of our axioms would be that for events E_1 and E_2 that are subsets of F , the ratio of the conditional probabilities $P(E_1|F)$ and $P(E_2|F)$ is the same as the ratio of $P(E_1)$ and $P(E_2)$.

she knows, that is, if she thinks that she knows the answer, then she really does.)
 What do you expect her grade to be on a 100 question True-False test to be?

For Exercise 6.3-2 let's let E be the event that the sum is even and F be the event that the sum is greater than or equal to 10. Thus referring to our sample space in Exercise 6.3-2, $P(F) = 1/6$ and $P(E \cap F) = 1/9$, since it is the probability that the roll is either 10 or 12. Dividing these two we get $2/3$.

In Exercise 6.3-3, the event that the total number of dots is odd has probability $1/2$. Similarly, given that the red die has an odd number of dots, the probability of an odd sum is $1/2$ since this event corresponds exactly to getting an even roll on the green die. Thus, by the definition of independence, the event of an odd number of dots on the red die and the event that the total number of dots is odd are independent.

In Exercise 6.3-4, if a student knows 80% of the material in a course, we would hope that her grade on a well-designed test of the course would be around 80%. But what if the test is a True-False test? Let R be the event that she gets the right answer, K be the event that she knows that right answer and \bar{K} be the event that she guesses. Then $R = P(R \cap K) + P(R \cap \bar{K})$. Since R is a union of two disjoint events, its probability would be the sum of the probabilities of these two events. How do we get the probabilities of these two events? The statement of the problem gives us implicitly the conditional probability that she gets the right answer given that she knows the answer, namely one, and the probability that she gets the right answer if she doesn't know the answer, namely $1/2$. Using Equation 6.14, we see that we use the equation

$$P(E \cap F) = P(E|F)P(F) \quad (6.15)$$

to compute $P(R \cap K)$ and $P(R \cap \bar{K})$, since the problem tells us directly that $P(K) = .8$ and $P(\bar{K}) = .2$. In symbols,

$$\begin{aligned} P(R) &= P(R \cap K) + P(R \cap \bar{K}) \\ &= P(R|K)P(K) + P(R|\bar{K})P(\bar{K}) \\ &= 1 \cdot .8 + .5 \cdot .2 = .9 \end{aligned}$$

We have shown that the probability that she gets the right answer is $.9$. Thus we would expect her to get a grade of 90%.

Independence

We said in Exercise 6.3-3 that E is independent of F if $P(E|F) = P(E)$. The *product principle for independent probabilities* (Theorem 6.4) gives another test for independence.

Theorem 6.4 *Suppose E and F are events in a sample space. Then E is independent of F if and only if $P(E \cap F) = P(E)P(F)$.*

Proof: First consider the case when F is non-empty. Then, from our definition in Exercise 6.3-3

$$E \text{ is independent of } F \quad \Leftrightarrow \quad P(E|F) = P(E).$$

(Even though the definition only has an “if”, recall the convention of using “if” in definitions, even though “if and only if” is meant.) Using the definition of $P(E|F)$ in Equation 6.14, in the right side of the above equation we get

$$\begin{aligned} P(E|F) &= P(E) \\ \Leftrightarrow \frac{P(E \cap F)}{P(F)} &= P(E) \\ \Leftrightarrow P(E \cap F) &= P(E)P(F). \end{aligned}$$

Since every step in this proof was an if and only if statement we have completed the proof for the case when F is non-empty.

If F is empty, then E is independent of F and both $P(E)P(F)$ and $P(E \cap F)$ are zero. Thus in this case as well, E is independent of F if and only if $P(E \cap F) = P(E)P(F)$. ■

Corollary 6.5 *E is independent of F if and only if F is independent of E .*

When we flip a coin twice, we think of the second outcome as being independent of the first. It would be a sorry state of affairs if our definition of independence did not capture this intuitive idea! Let’s compute the relevant probabilities to see if it does. For flipping a coin twice our sample space is $\{HH, HT, TH, TT\}$ and we weight each of these outcomes $1/4$. To say the second outcome is independent of the first, we must mean that getting an H second is independent of whether we get an H or a T first, and same for getting a T second. This gives us that $P(H \text{ first}) = 1/4 + 1/4 = 1/2$ and $P(H \text{ second}) = 1/2$, while $P(H \text{ first and H second}) = 1/4$. Note that

$$P(H \text{ first})P(H \text{ second}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = P(H \text{ first and H second}).$$

By Theorem 6.4, this means that the event “ H second” is independent of the event “ H first.” We can make a similar computation for each possible combination of outcomes for the first and second flip, and so we see that our definition of independence captures our intuitive idea of independence in this case. Clearly the same sort of computation applies to rolling dice as well.

Exercise 6.3-5 What sample space and probabilities have we been using when discussing hashing? Using these, show that the event “key i hashes to position p ” and the event “key j hashes to position q ” are independent when $i \neq j$. Are they independent if $i = j$?

In Exercise 6.3-5 if we have a list of n keys to hash into a table of size k , our sample space consists of all n -tuples of numbers between 1 and k . The event that key i hashes to some number p consists of all n -tuples with p in the i th position, so its probability is $\left(\frac{1}{k}\right)^{n-1} / \left(\frac{1}{k}\right)^n = \frac{1}{k}$. The probability that key j hashes to some number q is also $\frac{1}{k}$. If $i \neq j$, then the event that key i hashes to p and key j hashes to q has probability $\left(\frac{1}{k}\right)^{n-2} / \left(\frac{1}{k}\right)^n = \left(\frac{1}{k}\right)^2$, which is the product of the probabilities that key i hashes to p and key j hashes to q , so these two events are independent. However if $i = j$ the probability of key i hashing to p and key j hashing to q is zero unless $p = q$, in which case it is 1. Thus if $i = j$, these events are not independent.

Independent Trials Processes

Coin flipping and hashing are examples of what are called “independent trials processes.” Suppose we have a process that occurs in stages. (For example, we might flip a coin n times.) Let us use x_i to denote the outcome at stage i . (For flipping a coin n times, $x_i = H$ means that the outcome of the i th flip is a head.) We let S_i stand for the set of possible outcomes of stage i . (Thus if we flip a coin n times, $S_i = \{H, T\}$.) A process that occurs in stages is called an **independent trials process** if for each sequence a_1, a_2, \dots, a_n with $a_i \in S_i$,

$$P(x_i = a_i | x_1 = a_1, \dots, x_{i-1} = a_{i-1}) = P(x_i = a_i).$$

In other words, if we let E_i be the event that $x_i = a_i$, then

$$P(E_i | E_1 \cap E_2 \cap \dots \cap E_{i-1}) = P(E_i).$$

By our product principle for independent probabilities, this implies that

$$P(E_1 \cap E_2 \cap \dots \cap E_{i-1} \cap E_i) = P(E_1 \cap E_2 \cap \dots \cap E_{i-1})P(E_i). \quad (6.16)$$

Theorem 6.6 *In an independent trials process the probability of a sequence a_1, a_2, \dots, a_n of outcomes is $P(\{a_1\})P(\{a_2\}) \dots P(\{a_n\})$.*

Proof: We apply mathematical induction and Equation 6.16. ■

How do independent trials relate to coin flipping? Here our sample space consists of sequences of n H s and T s, and the event that we have an H (or T) on the i th flip is independent of the event that we have an H (or T) on each of the first $i - 1$ flips. In particular, the probability of an H on the i th flip is $2^{n-1}/2^n = .5$, and the probability of an H on the i th flip, given a particular sequence on the first $i - 1$ flips is $2^{n-i-1}/2^{n-i} = .5$.

How do independent trials relate to hashing a list of keys? As in Exercise 6.3-5 if we have a list of n keys to hash into a table of size k , our sample space consists of all n -tuples of numbers between 1 and k . The probability that key i hashes to p and keys 1 through $i - 1$ hash to q_1, q_2, \dots, q_{i-1} is $(\frac{1}{k})^{n-i} / (\frac{1}{k})^n$ and the probability that keys 1 through $i - 1$ hash to q_1, q_2, \dots, q_{i-1} is $(\frac{1}{k})^{n-i+1} / (\frac{1}{k})^n$. Therefore

$$P(\text{key } i \text{ hashes to } p | \text{keys 1 through } i - 1 \text{ hash to } q_1, q_2, \dots, q_{i-1}) = \frac{(\frac{1}{k})^{n-i} / (\frac{1}{k})^n}{(\frac{1}{k})^{n-i+1} / (\frac{1}{k})^n} = \frac{1}{k}.$$

Therefore, the event that key i hashes to some number p is independent of the event that the first $i - 1$ keys hash to some numbers q_1, q_2, \dots, q_{i-1} . Thus our model of hashing is an independent trials process.

Exercise 6.3-6 Suppose we draw a card from a standard deck of 52 cards, replace it, draw another card, and continue for a total of ten draws. Is this an independent trials process?

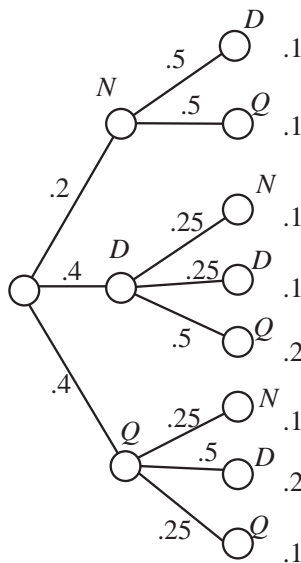
Exercise 6.3-7 Suppose we draw a card from a standard deck of 52 cards, discard it (i.e. we do not replace it), draw another card and continue for a total of ten draws. Is this an independent trials process?

In Exercise 6.3-6 we have an independent trials process, because the probability that we draw a given card at one stage does not depend on what cards we have drawn in earlier stages. However, in Exercise 6.3-7, we don't have an independent trials process. In the first draw, we have 52 cards to draw from, while in the second draw we have 51. In particular, we do not have the same cards to draw from on the second draw as the first, so the probabilities for each possible outcome on the second draw depend on whether that outcome was the result of the first draw.

Tree diagrams

When we have a sample space that consists of sequences of outcomes, it is often helpful to visualize the outcomes by a tree diagram. We will explain what we mean by giving a tree diagram of the following experiment. We have one nickel, two dimes, and two quarters in a cup. We draw a first and second coin. In Figure 6.3 you see our diagram for this process. Notice that in probability theory it is standard to have trees open to the right, rather than opening up or down.

Figure 6.5: A tree diagram illustrating a two-stage process.



Each level of the tree corresponds to one stage of the process of generating a sequence in our sample space. Each vertex is labeled by one of the possible outcomes at the stage it represents. Each edge is labeled with a conditional probability, the probability of getting the outcome at its right end given the sequence of outcomes that have occurred so far. Since no outcomes have occurred at stage 0, we label the edges from the root to the first stage vertices with the probabilities of the outcomes at the first stage. Each path from the root to the far right of the tree represents a possible sequence of outcomes of our process. We label each leaf node with the probability of the sequence that corresponds to the path from the root to that node. By the definition of conditional probabilities, the probability of a path is the product of the probabilities along its edges. We draw a probability tree for any (finite) sequence of successive trials in this way.

Sometimes a probability tree provides a very effective way of answering questions about a

process. For example, what is the probability of having a nickel in our coin experiment? We see there are four paths containing an N , and the sum of their weights is .4, so the probability that one of our two coins is a nickel is .4.

Exercise 6.3-8 How can we recognize from a probability tree whether it is the probability tree of an independent trials process?

Exercise 6.3-9 In Exercise 6.3-4 we asked (among other things), if a student knows 80% of the material in a course, what is the probability that she answers a question correctly on a 100 question True-False test (assuming that she guesses on any question she does not know the answer to)? (We assume that she knows what she knows, that is, if she thinks that she knows the answer, then she really does.) Show how we can use a probability tree to answer this question.

Exercise 6.3-10 A test for a disease that affects 0.1% of the population is 99% effective on people with the disease (that is, it says they have it with probability 0.99). The test gives a false reading (saying that a person who does not have the disease is affected with it) for 2% of the population without the disease. We can think of choosing someone and testing them for the disease as a two stage process. In stage 1, we either choose someone with the disease or we don't. In stage two, the test is either positive or it isn't. Give a probability tree for this process. What is the probability that someone selected at random and given a test for the disease will have a positive test? What is the probability that someone who has positive test results in fact has the disease?

A tree for an independent trials process has the property that at each level, for each node at that level, the (labeled) tree consisting of that node and all its children is identical to each labeled tree consisting of another node at that level and all its children. If we have such a tree, then it automatically satisfies the definition of an independent trials process.

In Exercise 6.3-9, if a student knows 80% of the material in a course, we expect that she has probability .8 of knowing the answer to any given question of a well-designed true-false test. We regard her work on a question as a two stage process; in stage 1 she determines whether she knows the answer, and in stage 2, she either answers correctly with probability 1, or she guesses, in which case she answers correctly with probability 1/2 and incorrectly with probability 1/2. Then as we see in Figure 6.3 there are two root-leaf paths corresponding to her getting a correct answer. One of these paths has probability .8 and the other has probability .1. Thus she actually has probability .9 of getting a right answer if she guesses at each question she does not know the answer to.

In Figure 6.3 we show the tree diagram for thinking of Exercise 6.3-10 as a two stage process. In the first stage, a person either has or doesn't have the disease. In the second stage we administer the test, and its result is either positive or not. We use D to stand for having the disease and ND to stand for not having the disease. We use "pos" to stand for a positive test and "neg" to stand for a negative test, and assume a test is either positive or negative. The question asks us for the conditional probability that someone has the disease, given that they test positive. This is

$$P(D|\text{pos}) = \frac{P(D \cap \text{pos})}{P(\text{pos})}.$$

Figure 6.6: The probability of getting a right answer is .9.

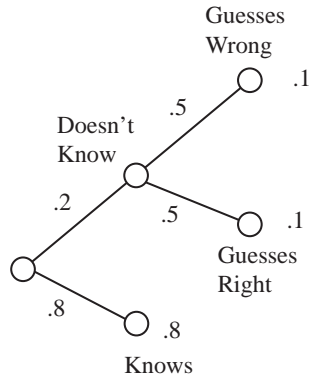
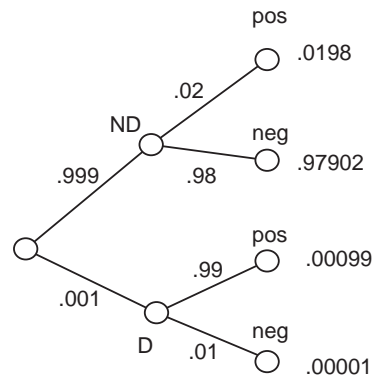


Figure 6.7: A tree diagram illustrating Exercise 6.3-10.



From the tree, we read that $P(D \cap \text{pos}) = .00099$ because this event consists of just one root-leaf paths. The event “pos” consists of two root-leaf paths whose probabilities total $.0198 + .00099 = .02097$. Thus $P(D|\text{pos}) = P(D \cap \text{pos})/P(\text{pos}) = .00099/.02097 = .0472$. Thus, given a disease this rare and a test with this error rate, a positive result only gives you roughly a 5% chance of having the disease! Here is another instance where a probability analysis shows something we might not have expected initially. This explains why doctors often don’t want to administer a test to someone unless that person is already showing some symptoms of the disease being tested for.

We can also do Exercise 6.3-10 purely algebraically. We are given that

$$P(\text{disease}) = .001, \quad (6.17)$$

$$P(\text{positive test result}|\text{disease}) = .99, \quad (6.18)$$

$$P(\text{positive test result}|\text{no disease}) = .02. \quad (6.19)$$

We wish to compute

$$P(\text{disease}|\text{positive test result}).$$

We use Equation 6.14 to write that

$$P(\text{disease}|\text{positive test result}) = \frac{P(\text{disease} \cap \text{positive test result})}{P(\text{positive test result})}. \quad (6.20)$$

How do we compute the numerator? Using the fact that $P(\text{disease} \cap \text{positive test result}) = P(\text{positive test result} \cap \text{disease})$ and Equation 6.14 again, we can write

$$P(\text{positive test result}|\text{disease}) = \frac{P(\text{positive test result} \cap \text{disease})}{P(\text{disease})}.$$

Plugging Equations 6.18 and 6.17 into this equation, we get

$$.99 = \frac{P(\text{positive test result} \cap \text{disease})}{.001}$$

or $P(\text{positive test result} \cap \text{disease}) = (.001)(.99) = .00099$.

To compute the denominator of Equation 6.20, we observe that since each person either has the disease or doesn't, we can write

$$P(\text{positive test}) = P(\text{positive test} \cap \text{disease}) + P(\text{positive test} \cap \text{no disease}). \quad (6.21)$$

We have already computed $P(\text{positive test result} \cap \text{disease})$, and we can compute the probability $P(\text{positive test result} \cap \text{no disease})$ in a similar manner. Writing

$$P(\text{positive test result}|\text{no disease}) = \frac{P(\text{positive test result} \cap \text{no disease})}{P(\text{no disease})},$$

observing that $P(\text{no disease}) = 1 - P(\text{disease})$ and plugging in the values from Equations 6.17 and 6.19, we get that $P(\text{positive test result} \cap \text{no disease}) = (.02)(1 - .001) = .01998$. We now have the two components of the right hand side of Equation 6.21 and thus $P(\text{positive test result}) = .00099 + .01998 = .02097$. Finally, we have all the pieces in Equation 6.20, and conclude that

$$P(\text{disease}|\text{positive test result}) = \frac{P(\text{disease} \cap \text{positive test result})}{P(\text{positive test result})} = \frac{.00099}{.02097} = .0472.$$

Clearly, using the tree diagram mirrors these computations, but it both simplifies the thought process and reduces the amount we have to write.

Important Concepts, Formulas, and Theorems

1. *Conditional Probability.* We define the *conditional probability* of E given F , denoted by $P(E|F)$ and read as "the probability of E given F " by

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$

2. *Independent.* We say E is *independent* of F if $P(E|F) = P(E)$.
3. *Product Principle for Independent Probabilities.* The *product principle for independent probabilities* (Theorem 6.4) gives another test for independence. Suppose E and F are events in a sample space. Then E is independent of F if and only if $P(E \cap F) = P(E)P(F)$.

4. *Symmetry of Independence.* The event E is independent of the event F if and only if F is independent of E .
5. *Independent Trials Process.* A process that occurs in stages is called an *independent trials process* if for each sequence a_1, a_2, \dots, a_n with $a_i \in S_i$,

$$P(x_i = a_i | x_1 = a_1, \dots, x_{i-1} = a_{i-1}) = P(x_i = a_i).$$

6. *Probabilities of Outcomes in Independent Trials.* In an independent trials process the probability of a sequence a_1, a_2, \dots, a_n of outcomes is $P(\{a_1\})P(\{a_2\}) \cdots P(\{a_n\})$.
7. *Coin Flipping.* Repeatedly flipping a coin is an independent trials process.
8. *Hashing.* Hashing a list of n keys into k slots is an independent trials process with n stages.
9. *Probability Tree.* In a probability tree for a multistage process, each level of the tree corresponds to one stage of the process. Each vertex is labeled by one of the possible outcomes at the stage it represents. Each edge is labeled with a conditional probability, the probability of getting the outcome at its right end given the sequence of outcomes that have occurred so far. Each path from the root to a leaf represents a sequence of outcomes and is labelled with the product of the probabilities along that path. This is the probability of that sequence of outcomes.

Problems

1. In three flips of a coin, what is the probability that two flips in a row are heads, given that there is an even number of heads?
2. In three flips of a coin, is the event that two flips in a row are heads independent of the event that there is an even number of heads?
3. In three flips of a coin, is the event that we have at most one tail independent of the event that not all flips are identical?
4. What is the sample space that we use for rolling two dice, a first one and then a second one? Using this sample space, explain why it is that if we roll two dice, the event “ i dots are on top of the first die” and the event “ j dots are on top of the second die” are independent.
5. If we flip a coin twice, is the event of having an odd number of heads independent of the event that the first flip comes up heads? Is it independent of the event that the second flip comes up heads? Would you say that the three events are mutually independent? (This hasn’t been defined, so the question is one of opinion. However you should back up your opinion with a reason that makes sense!)
6. Assume that on a true-false test, students will answer correctly any question on a subject they know. Assume students guess at answers they do not know. For students who know 60% of the material in a course, what is the probability that they will answer a question correctly? What is the probability that they will know the answer to a question they answer correctly?

7. A nickel, two dimes, and two quarters are in a cup. We draw three coins, one at a time, without replacement. Draw the probability tree which represents the process. Use the tree to determine the probability of getting a nickel on the last draw. Use the tree to determine the probability that the first coin is a quarter, given that the last coin is a quarter.
8. Write down a formula for the probability that a bridge hand (which is 13 cards, chosen from an ordinary deck) has four aces, given that it has one ace. Write down a formula for the probability that a bridge hand has four aces, given that it has the ace of spades. Which of these probabilities is larger?
9. A nickel, two dimes, and three quarters are in a cup. We draw three coins, one at a time without replacement. What is the probability that the first coin is a nickel? What is the probability that the second coin is a nickel? What is the probability that the third coin is a nickel?
10. If a student knows 75% of the material in a course, and a 100 question multiple choice test with five choices per question covers the material in a balanced way, what is the student's probability of getting a right answer to a given question, given that the student guesses at the answer to each question whose answer he or she does not know?
11. Suppose E and F are events with $E \cap F = \emptyset$. Describe when E and F are independent and explain why.
12. What is the probability that in a family consisting of a mother, father and two children of different ages, that the family has two girls, given that one of the children is a girl? What is the probability that the children are both boys, given that the older child is a boy?