# PROBLEM SET 1

## Basic Probability Concepts

1. A survey of 1000 people determines that 80% like walking and 60% like biking, and all like at least one of the two activities. What is the probability that a randomly chosen person in this survey likes biking but not walking?

   A) 0 B) 0.1 C) 0.2 D) 0.3 E) 0.4

2. (SOA) Among a large group of patients recovering from shoulder injuries, it is found that 22% visit both a physical therapist and a chiropractor, whereas 12% visit neither of these. The probability that a patient visits a chiropractor exceeds by 0.14 the probability that a patient visits a physical therapist. Determine the probability that a randomly chosen member of this group visits a physical therapist.

   A) 0.26 B) 0.38 C) 0.40 D) 0.48 E) 0.62

3. (SOA) An insurer offers a health plan to the employees of a large company. As part of this plan, the individual employees may choose exactly two of the supplementary coverages A, B, and C, or they may choose no supplementary coverage. The proportions of the company's employees that choose suuplementary coverages A, B, and C are $\frac{1}{4}$, $\frac{1}{3}$, and $\frac{5}{12}$, respectively. Determine the probability that a randomly chosen employee will choose no supplementary coverage.

   A) 0 B) $\frac{47}{144}$ C) $\frac{1}{2}$ D) $\frac{97}{144}$ E) $\frac{7}{9}$

4. (SOA) An auto insurance company has 10,000 policyholders.
   Each policyholder is classified as
   (i) young or old;
   (ii) male or female; and
   (iii) married or single.
   Of these policyholders, 3000 are young, 4600 are male, and 7000 are married. The policyholders can also be classified as 1320 young males, 3010 married males, and 1400 young married persons. Finally, 600 of the policyholders are young married males. How many of the company's policyholders are young, female, and single?

   A) 280 B) 423 C) 486 D) 880 E) 896

5. (SOA) The probability that a visit to a primary care physician's (PCP) office results in neither lab work nor referral to a specialist is 35%. Of those coming to a PCP's office, 30% are referred to specialists and 40% require lab work. Determine the probability that a visit to a PCP's office results in both lab work and referral to a specialist.

   A) 0.05 B) 0.12 C) 0.18 D) 0.25 E) 0.35

6. (SOA) You are given $P[A \cup B] = 0.7$ and $P[A \cup B'] = 0.9$. Determine $P[A]$.
A) 0.2 B) 0.3 C) 0.4 D) 0.6 E) 0.8

7. (SOA) A survey of a group's viewing habits over the last year revealed the following information:

(i) 28% watched gymnastics
(ii) 29% watched baseball
(iii) 19% watched soccer
(iv) 14% watched gymnastics and baseball
(v) 12% watched baseball and soccer
(vi) 10% watched gymnastics and soccer
(vii) 8% watched all three sports.

Calculate the percentage of the group that watched none of the three sports during the last year.
A) 24 B) 36 C) 41 D) 52 E) 60

8. (SOA) Under an insurance policy, a maximum of five claims may be filed per year by a policyholder. Let $p_n$ be the probability that a policyholder files $n$ claims during a given year, where $n = 0, 1, 2, 3, 4, 5$ . An actuary makes the following observations:
i) $p_n \geq p_{n+1}$ for $n = 0, 1, 2, 3, 4$.
ii) The difference between $p_n$ and $p_{n+1}$ is the same for $n = 0, 1, 2, 3, 4$.
iii) Exactly 40% of policyholders file fewer than two claims during a given year.
Calculate the probability that a random policyholder will file more than three claims during a given year.
A) 0.14 B) 0.16 C) 0.27 D) 0.29 E) 0.33

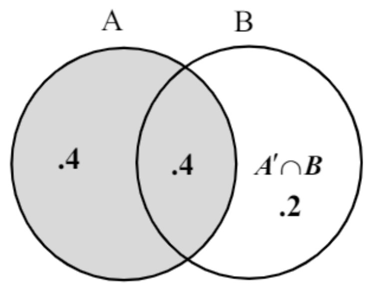
9. (SOA) The probability that a member of a certain class of homeowners with liability and property coverage will file a liability claim is 0.04, and the probability that a member of this class will file a property claim is 0.10. The probability that member of this class will file a liability claim but not a property claim is 0.01. Calculate the probability that a randomly selected member of this class of homeowners will not file a claim of either type.
A) 0.850 B) 0.860 C) 0.864 D) 0.870 E) 0.890

10. (SOA) A mattress store sells only, king, queen and twin-size mattresses. Sales records at the store indicate that one-fourth as many queen-size mattresses are sold as king and twin-size mattresses combined. Records also indicate that three times as many king-size mattresses are sold as twin-size mattresses.
Calculate the probability that the next mattress sold is either king or queen-size
A) 0.12 B) 0.13 C) 0.80 D) 0.85 E) 0.95

# PROBLEM SET 1 SOLUTIONS

1. Let $A =$ "like walking" and $B =$ "like biking". We use the interpretation that "percentage" and "proportion" are taken to mean "probability".

We are given $P(A) = 0.8$, $P(B) = 0.6$ and $P(A \cup B) = 1$.
From the diagram below we can see that since $A \cup B = A \cup (B \cap A')$ we have
$P(A \cup B) = P(A) + P(A' \cap B) \rightarrow P(A' \cap B) = 0.2$ is the proportion of people who like biking but (and) not walking. In a similar way we get $P(A \cap B') = 0.4$.




An algebraic approach is the following. Using the rule $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, we get $1 = .8 + .6 - P(A \cap B) \rightarrow P(A \cap B) = 0.4$. Then, using the rule
$P(B) = P(B \cap A) + P(B \cap A')$, we get $P(B \cap A') = 0.6 - 0.4 = 0.2$. Answer: C

2. $C$ - chiropractor visit ; $T$ - therapist visit.
We are given $P(C \cap T) = 0.22$ , $P(C' \cap T') = 0.12$ , $P(C) = P(T) + 0.14$.
$0.88 = 1 - P(C' \cap T') = P(C \cup T) = P(C) + P(T) - P(C \cap T)$
$= P(T) + 0.14 + P(T) - 0.22 \rightarrow P(T) = 0.48$. Answer: D

3. Since someone who chooses coverage must choose exactly two supplementary coverages, in order for someone to choose coverage A, they must choose either A-and-B or A-and-C. Thus, the proportion of $\frac{1}{4}$ of individuals that choose A is $P[A \cap B] + P[A \cap C] = \frac{1}{4}$ (where this refers to the probability that someone chosen at random in the company chooses coverage A). Similarly, $P[B \cap A] + P[B \cap C] = \frac{1}{3}$ and $P[C \cap A] + P[C \cap B] = \frac{5}{12}$. Then,
$$(P[A \cap B] + P[A \cap C]) + (P[B \cap A] + P[B \cap C]) + (P[C \cap A] + P[C \cap B])$$
$$= 2(P[A \cap B] + P[A \cap C] + P[B \cap C]) = \frac{1}{4} + \frac{1}{3} + \frac{5}{12} = 1.$$
It follows that $P[A \cap B] + P[A \cap C] + P[B \cap C] = \frac{1}{2}$.
This is the probability that a randomly chosen individual chooses some form of supplementary coverage, since if someone who chooses coverage chooses exactly two of A, B, and C. Therefore, the probability that a randomly chosen individual does not choose any coverage is the probability of the complementary event, which is also $\frac{1}{2}$. Answer: C

4. We identify the following subsets of the set of 10,000 policyholders:

$Y =$ young, with size 3000 (so that $Y' =$ old has size 7000),

$M =$ male, with size 4600 (so that $M' =$ female has size 5400), and

$C =$ married, with size 7000 (so that $C' =$ single has size 3000).

We are also given that $Y \cap M$ has size 1320 , $M \cap C$ has size 3010,

$Y \cap C$ has size 1400 , and $Y \cap M \cap C$ has size 600.

We wish to find the size of the subset $Y \cap M' \cap C'$.

We use the following rules of set theory:

(i) if two finite sets are disjoint (have no elements in common, also referred to as empty intersection), then the total number of elements in the union of the two sets is the sum of the numbers of elements in each of the sets;

(ii) for any sets $A$ and $B$, $A = (A \cap B) \cup (A \cap B')$ , and $A \cap B$ and $A \cap \overline{B}$ are disjoint.

Applying rule (ii), we have $Y = (Y \cap M) \cup (Y \cap M')$. Applying rule (i), it follows that the size of $Y \cap M'$ must be $3000 - 1320 = 1680$.
We now apply rule (ii) to $Y \cap C$ to get $Y \cap C = (Y \cap C \cap M) \cup (Y \cap C \cap M')$ .
Applyng rule (i), it follows that $Y \cap C \cap M'$ has size $1400 - 600 = 800$.

Now applying rule (ii) to $Y \cap M'$ we get $Y \cap M' = (Y \cap M' \cap C) \cup (Y \cap M' \cap C')$.
Applying rule (i), it follows that $Y \cap M' \cap C'$ has size $1680 - 800 = 880$.

Within the "Young" category, which we are told is 3000, we can summarize the calculations in the following table. This is a more straightforward solution.

| | Married<br>1400 (given) | Single<br>$1600 = 3000 - 1400$ |
|---|---|---|
| Male<br>1320 (given) | 600 (given) | $720 = 1320 - 600$ |
| Female<br>$1680 =$<br>$3000 - 1320$ | $800 = 1400 - 600$ | $880 = 1600 - 720$ |

Answer: D

5. We identify events as follows:

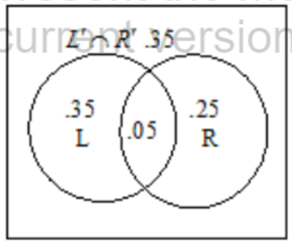
$L$ : lab work needed

$R$ : referral to a specialist needed

We are given $P[L' \cap R'] = 0.35$ , $P[R] = 0.3$ , $P[L] = 0.4$. It follows that

$P[L \cup R] = 1 - P[L' \cap R'] = 0.65$ , and then since

$P[L \cup R] = P[L] + P[R] - P[L \cap R]$, we get $P[L \cap R] = 0.3 + 0.4 - 0.65 = 0.05$ .




These calculations can be summarized in the following table

| | | |
|---|---|---|
| | $L$ , 0.4<br>given | $L'$ , 0.6<br>$0.6 = 1 - 0.4$ |
| $R$ , 0.3<br>given | $L \cap R$<br>$0.05 = 0.4 - 0.35$ | $L' \cap R$<br>$0.25 = 0.3 - 0.05$ |
| $R'$ , 0.7<br>$0.7 = 1 - 0.3$ | $L \cap R'$<br>$0.35 = 0.7 - 0.35$ | $L' \cap R'$ , 0.35<br>given |

Answer: A

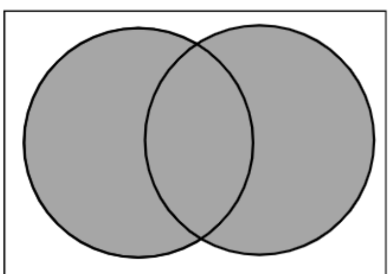
6. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$ , $P[A \cup B'] = P[A] + P[B'] - P[A \cap B']$.

We use the relationship $P[A] = P[A \cap B] + P[A \cap B']$. Then

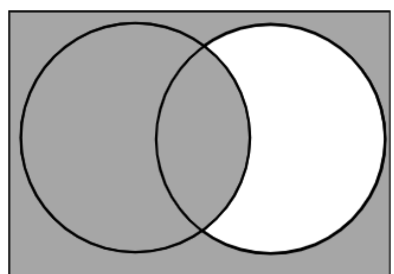
$$P[A \cup B] + P[A \cup B'] = P[A] + P[B] - P[A \cap B] + P[A] + P[B'] - P[A \cap B']$$
$$= 2P[A] + 1 - P[A] = P[A] + 1 \text{ (since } P[B] + P[B'] = 1).$$

Therefore, $0.7 + 0.9 = P[A] + 1$ so that $P[A] = 0.6$.

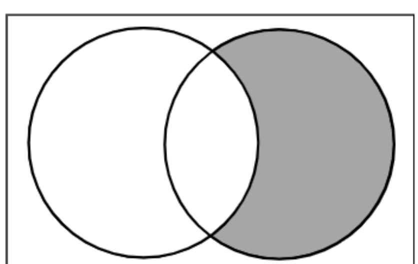
An alternative solution is based on the following Venn diagrams.



$P[A \cup B] = .70$



$P[A \cup B'] = .90$



$P[(A \cup B')'] = P[A' \cap B] = .10$

In the third diagram, the shaded area is the complement of that in the second diagram (using De Morgan's Law, we have $(A \cup B')' = A' \cap B'' = A' \cap B$ ). Then it can be seen from diagrams 1 and 3 that $A = (A \cup B) - (A' \cap B)$, so that $P[A] = P[A \cup B] - P[A' \cap B] = 0.7 - 0.1 = 0.6.$ Answer: D

7. We identify the following events:
$G$ - watched gymnastics , $B$ - watched baseball , $S$ - watched soccer .
We wish to find $P[G' \cap B' \cap S']$. By DeMorgan's rules we have
$P[G' \cap B' \cap S'] = 1 - P[G \cup B \cup S]$.
We use the relationship

$$P[G \cup B \cup S] = P[G] + P[B] + P[S] - \left(P[G \cap B] + P[G \cap S] + P[B \cap S]\right) + P[G \cap B \cap S].$$

We are given $P[G] = 0.28$ , $P[B] = 0.29$ , $P[S] = 0.19$,
$P[G \cap B] = 0.14, P[G \cap S] = 0.10, P[B \cap S] = 0.12,\ P[G \cap B \cap S] = 0.08.$
Then $P[G \cup B \cup S] = 0.48$ and $P[G' \cap B' \cap S'] = 1 - 0.48 = 0.52.$ Answer: D

8. The probability in question is $p_4 + p_5$. We know that $p_0 + p_1 + p_2 + p_3 + p_4 + p_5 = 1$.
We are given that $p_0 - p_1 = p_1 - p_2 = p_2 - p_3 = p_3 - p_4 = p_4 - p_5$,
and $p_0 + p_1 = .4$ . If we let $t$ be equal to the common difference $p_{n+1} - p_n$ , then
$p_1 = p_0 + t$ , $p_2 = p_0 + 2t$ , $p_3 = p_0 + 3t$ , $p_4 = p_0 + 4t$ and $p_5 = p_0 + 5t$ .
Then $p_0 + p_1 + p_2 + p_3 + p_4 + p_5 = 6p_0 + 15t = 1$ . When we combine this equation with $p_0 + p_1 = 2p_0 + t = .4$ , we can solve the two equations to get $p_0 = \frac{5}{24}$ , and $t = -\frac{1}{60}$ .
Then $p_4 = p_0 + 4t = \frac{17}{120}$ , and $p_5 = p_0 + 5t = \frac{15}{120}$ , so that $p_4 + p_5 = \frac{32}{120} = 0.27.$ Answer: C

9. We define the following events: $L =$ file a liability claim , $P =$ file a property claim.
We are given $P(L) = 0.04$ , $P(P) = 0.10$ , $P(L \cap P') = 0.01$. We wish to find $P(L' \cap P')$.
$P(P') = 1 - P(P) = 0.90$ and $0.90 = P(P') = P(L \cap P') + P(L' \cap P') = 0.01 + P(L' \cap P')$.
It follows that $P(L' \cap P') = 0.89.$ Answer: E

10. We define $T$ to be the event that the next mattress sold is twin-size, and similarly we define $K$ and $Q$ as the events that the next mattress sold is king-size and queen-size, respectively. We define $P(T) = c$. Then $P(K) = 3c$ and $P(Q) = \frac{1}{4} \times [P(K) + P(T)] = c$.
Since $P(T) + P(Q) + P(K) = 1$, we have $5c = 1$ , so that $c = 0.2$.
Then $P(K \cup Q) = 4c = 0.8.$ Answer: C

# SECTION 2 - CONDITIONAL PROBABILITY AND INDEPENDENCE

## Conditional probability of event $B$ given event $A$:

If $P[A] > 0$, then $\boldsymbol{P[B|A] = \frac{P[B \cap A]}{P[A]}}$ is defined to be the conditional probability that event $B$ occurs given that event $A$ has occurred. Events $A$ and $B$ may be related so that if we know that event $A$ has occurred, the **conditional probability of event $B$ occurring given that event $A$ has occurred** might not be the same as the unconditional probability of event $B$ occurring if we had no knowledge about the occurrence of event $A$. For instance, if a fair 6-sided die is tossed and if we know that the outcome is even, then the conditional probability is 0 of tossing a 3 given that the toss is even. If we did not know that the toss was even, if we had no knowledge of the nature of the toss, then tossing a 3 would have an unconditional probability of $\frac{1}{6}$, the same as all other possible tosses that could occur.

When we condition on event $A$, we are assuming that event $A$ has occurred so that $A$ becomes the new probability space, and all conditional events must take place within event $A$ (the new probability space). Dividing by $P[A]$ scales all probabilities so that $A$ is the entire probability space, and $P[A|A] = 1$. To say that event $B$ has occurred given that event $A$ has occurred means that both $B$ and $A$ ($B \cap A$) have occurred within the probability space $A$. This explains the numerator $P(B \cap A)$ in the definition of the conditional probability $P[B|A]$.

Rewriting $P[B|A] = \frac{P[B \cap A]}{P[A]}$, the equation that defines conditional probability, results in $\boldsymbol{P[B \cap A] = P[B|A] \cdot P[A]}$, which is referred to as the **multiplication rule**.

**Example 2-1:** Suppose that a fair six-sided die is tossed. The probability space is $S = \{1, 2, 3, 4, 5, 6\}$. We define the following events:
$A =$ "the number tossed is even" $= \{2, 4, 6\}$, $B =$ "the number tossed is $\leq 3$" $= \{1, 2, 3\}$,
$C =$ "the number tossed is a 1 or a 2" $= \{1, 2\}$,
$D =$ "the number tossed doesn't start with the letters 'f' or 't'" $= \{1, 6\}$.

The conditional probability of $B$ given $A$ is

$P[B|A] = \frac{P[\{1,2,3\} \cap \{2,4,6\}]}{P[\{2,4,6\}]} = \frac{P[\{2\}]}{P[\{2,4,6\}]} = \frac{1/6}{1/2} = \frac{1}{3}$. The interpretation of this conditional probability is that if we know that event $A$ has occurred, then the toss must be 2, 4 or 6. Since the original 6 possible tosses of a die were equally likely, if we are given the additional information that the toss is 2, 4 or 6, it seems reasonable that each of those is equally likely, each with a probability of $\frac{1}{3}$. Then within the reduced probability space $A$, the (conditional) probability that event $B$ occurs is the probability, in the reduced space, of tossing a 2; this is $\frac{1}{3}$.

For events $B$ and $C$ defined above, the conditional probability of $B$ given $C$ is $P[B|C] = 1$.
To say that $C$ has occurred means that the toss is 1 or 2. It is then guaranteed that event $B$ has occurred (the toss is a 1, 2 or 3), since $C \subset B$.
The conditional probability of $A$ given $C$ is $P[A|C] = \frac{1}{2}$. □

**Example 2-2:** If $P[A] = \frac{1}{6}$ and $P[B] = \frac{5}{12}$, and $P[A|B] + P[B|A] = \frac{7}{10}$, find $P[A \cap B]$.
**Solution:** $P[B|A] = \frac{P[A \cap B]}{P[A]} = 6P[A \cap B]$ and $P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{12}{5} P[A \cap B]$
$\rightarrow (6 + \frac{12}{5}) \times P[A \cap B] = \frac{7}{10} \rightarrow P[A \cap B] = \frac{1}{12}$. □

**IMPORTANT NOTE:** The following manipulation of event probabilities arises from time to time:

$$P[B] = P[B|A] \times P(A) + P[B|A'] \times P(A').$$

This relationship is a version of the **Law of Total Probability.** This relationship is valid since for any events $A$ and $B$, we have $P[B] = P[B \cap A] + P[B \cap A']$. We then use the relationships $P[B \cap A] = P[B|A] \times P(A)$ and $P[B \cap A'] = P[B|A'] \times P(A')$. If we know the conditional probabilities for event $B$ given some other event $A$ and if we also know the conditional probability of $B$ given the complement $A'$, and if we are given the (unconditional) probability of event $A$, then we can find the (unconditional) probability of event $B$. An application of this concept occurs when an experiment has two (or more) steps. The following example illustrates this idea.

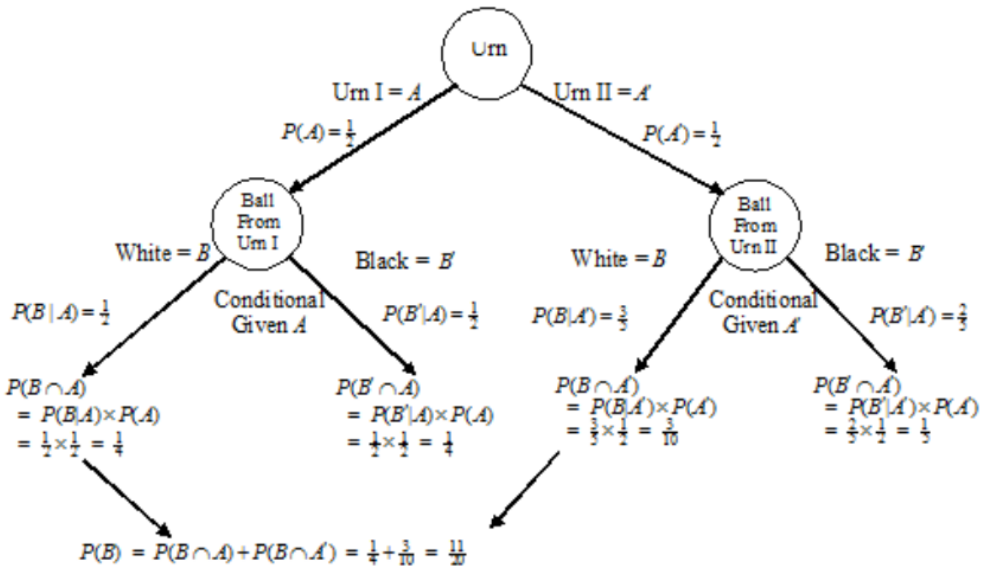
**Example 2-3:**
Urn I contains 2 white and 2 black balls and Urn II contains 3 white and 2 black balls. An Urn is chosen at random, and a ball is randomly selected from that Urn. Find the probability that the ball chosen is white.
**Solution:**
Let $A$ be the event that Urn I is chosen and $A'$ is the event that Urn II is chosen. The implicit assumption is that both Urns are equally likely to be chosen (this is the meaning of "an Urn is chosen at random"). Therefore, $P[A] = \frac{1}{2}$ and $P[A'] = \frac{1}{2}$. Let $B$ be the event that the ball chosen is white. If we know that Urn I was chosen, then there is $\frac{1}{2}$ probability of choosing a white ball (2 white out of 4 balls, it is assumed that each ball has the same chance of being chosen); this can be described as $P[B|A] = \frac{1}{2}$. In a similar way, if Urn II is chosen, then $P[B|A'] = \frac{3}{5}$ (3 white out of 5 balls). We can now apply the relationship described prior to this example.
$P[B \cap A] = P[B|A] \times P[A] = (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4}$, and $P[B \cap A'] = P[B|A'] \times P[A'] = \frac{3}{5} \times \frac{1}{2} = \frac{3}{10}$. Finally, $P[B] = P[B \cap A] + P[B \cap A'] = \frac{1}{4} + \frac{3}{10} = \frac{11}{20}$.

The order of calculations (1-2-3) can be summarized in the following table

| | $A$ | $A'$ |
|---|---|---|
| $B$ | **1.** $P(B\cap A) = P[B\vert A]\times P[A]$ | **2.** $P(B\cap A') = P[B\vert A']\times P[A']$ |

**3.** $P(B) = P(B\cap A) + P(B\cap A')$

An event tree diagram, shown below, is another way of illustrating the probability relationships.




□

**IMPORTANT NOTE:** An exam question may state that "an item is to be chosen at random from a collection of items". Unless there is an indication otherwise, this is interpreted to mean that each item has the same chance of being chosen. Also, if we are told that a fair coin is tossed randomly, then we interpret that to mean that the head and tail each have the probability of 0.5 occurring. Of course, if we are told that the coin is "loaded" so that the probability of tossing a head is 2/3 and tail is 1/3, then random toss means the head and tail will occur with those stated probabilities.

## Bayes' rule and Bayes' Theorem (basic form):

For any events $A$ and $B$ with $P[B]>0$, $\boldsymbol{P[A|B]=\frac{P[A\cap B]}{P[B]}} = \boldsymbol{\frac{P[B|A]\times P[A]}{P[B|A]\times P[A]+P[B|A']\times P[A']}}$ .

The usual way that this is applied is in the case that we are given the values of
$P[A]$ , $P[B|A]$ and $P[B|A']$ (from $P[A]$ we get $P[A'] = 1-P[A]$),
and we are asked to find $P[A|B]$ (in other words, we are asked to "turn around" the conditioning of the events $A$ and $B$). We can summarize this process by calculating the quantities in the following table in the order indicated numerically (1-2-3-4) (other entries in the table are not necessary in this calculation, but might be needed in related calculations).

| | $A$ , $P(A)$ given | $A'$ , $P(A')$ given |
|---|---|---|
| $B$ | $P[B\|A]$ given<br>**1.** $P[B \cap A] = P[B\|A] \times P[A]$ | $P[B\|A']$ given<br>**2.** $P[B \cap A'] = P[B\|A'] \times P[A']$ |

$$\Downarrow \qquad\qquad \Downarrow$$

**3.** $P[B] = P[B \cap A] + P[B \cap A']$

Also, we can find

| | | |
|---|---|---|
| $B'$ | $P[B'\|A] = 1 - P[B\|A]$<br>$P[B' \cap A] = P[B'\|A] \times P[A]$ | $P[B'\|A'] = 1 - P[B\|A']$<br>$P[A' \cap B'] = P[B'\|A'] \times P[A']$ |

and $P[B'] = P[B' \cap A] + P[B' \cap A']$

(but we could have found $P[B']$ from $P[B'] = 1 - P[B]$ , once $P[B]$ was found).

Step **4**: $P[A|B] = \frac{P[A \cap B]}{P[B]}$

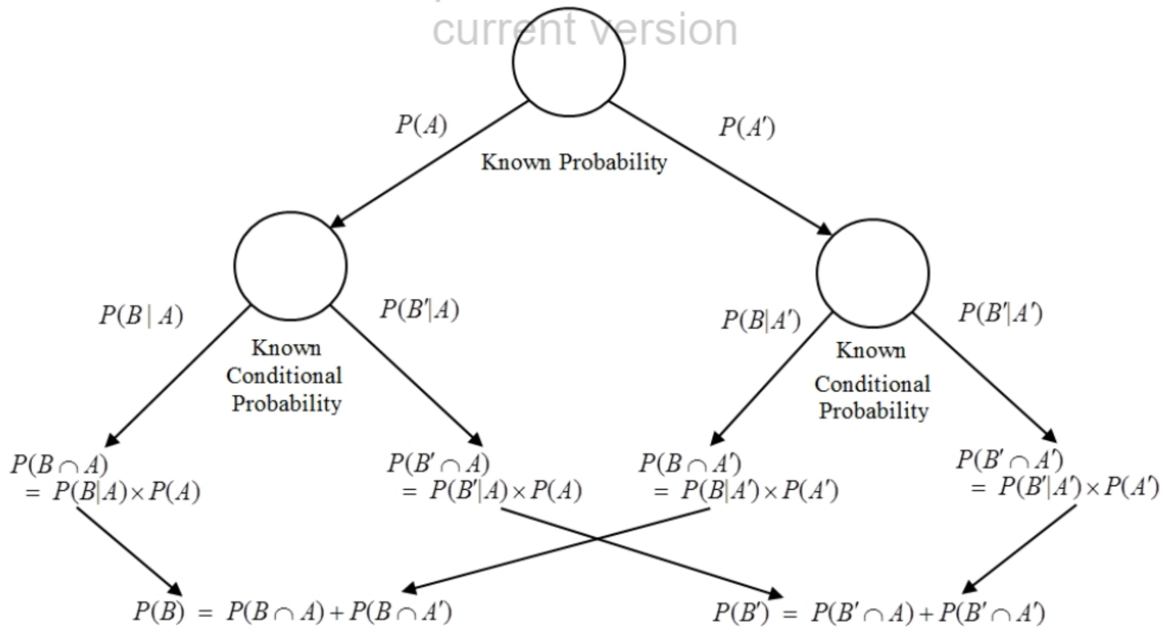
This can also be summarized in the following sequence of calculations.

$$\begin{array}{ccc} P[A]\,,\ P[B|A]\,,\ \text{given} & & P[A'] = 1 - P[A]\,,\ P[B|A']\,,\ \text{given} \\ \Downarrow & & \Downarrow \\ P[B \cap A] = P[B|A] \times P[A] & & P[B \cap A'] = P[B|A'] \times P[A'] \end{array} \quad \Downarrow$$

$$P[B] = P[B \cap A] + P[B \cap A']$$

Algebraically, we have done the following calculation:

$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[A \cap B]}{P[B \cap A] + P[B \cap A']} = \frac{P[B|A] \times P[A]}{P[B|A] \times P[A] + P[B|A'] \times P[A']}$,
where all the factors in the final expression were originally known. Note that the numerator is one of the components of the denominator. The following event tree is similar to the one in Example 2-3, illustrating the probability relationships.




Note that at the bottom of the event tree, $P(B')$ is also equal to $1 - P(B)$.

**Exam questions that involve conditional probability and make use of Bayes rule (and its extended form reviewed below) occur frequently. The key starting point is identifying and labeling unconditional events and conditional events and their probabilities in an efficient way.**

**Example 2-4:**

Urn I contains 2 white and 2 black balls and Urn II contains 3 white and 2 black balls. One ball is chosen at random from Urn I and transferred to Urn II, and then a ball is chosen at random from Urn II. The ball chosen from Urn II is observed to be white. Find the probability that the ball transferred from Urn I to Urn II was white.

**Solution:**

Let $A$ denote the event that the ball transferred from Urn I to Urn II was white and let $B$ denote the event that the ball chosen from Urn II is white. We are asked to find $P[A|B]$.

From the simple nature of the situation (and the usual assumption of uniformity in such a situation, meaning that all balls are equally likely to be chosen from Urn I in the first step), we have $P[A] = \frac{1}{2}$ (2 of the 4 balls in Urn I are white), and $P[A'] = \frac{1}{2}$.

If the ball transferred is white, then Urn II has 4 white and 2 black balls, and the probability of choosing a white ball out of Urn II is $\frac{2}{3}$; this is $P[B|A] = \frac{2}{3}$.

For trial use only.
Content may not
represent the most
current version

If the ball transferred is black, then Urn II has 3 white and 3 black balls, and the probability of choosing a white ball out of Urn II is $\frac{1}{2}$; this is $P[B|A'] = \frac{1}{2}$.

All of the information needed has been identified. From the table described above, we do the calculations in the following order:

1. $P[B \cap A] = P[B|A] \times P[A] = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$
2. $P[B \cap A'] = P[B|A'] \times P[A'] = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
3. $P[B] = P[B \cap A] + P[B \cap A'] = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$
4. $P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{1/3}{7/12} = \frac{4}{7}$ □

**Example 2-5:**

Identical twins come from the same egg and hence are of the same sex. Fraternal twins have a 50-50 chance of being the same sex. Among twins, the probability of a fraternal set is $p$ and an identical set is $q = 1 - p$. If the next set of twins are of the same sex, what is the probability that they are identical?

**Solution:**

Let $B$ be the event "the next set of twins are of the same sex", and let $A$ be the event "the next sets of twins are identical". We are given $P[B|A] = 1$, $P[B|A'] = 0.5$, $P[A] = q$, $P[A'] = p = 1 - q$.
Then $P[A|B] = \frac{P[B \cap A]}{P[B]}$ is the probability we are asked to find.
But $P[B \cap A] = P[B|A] \times P[A] = q$, and $P[B \cap A'] = P[B|A'] \times P[A'] = 0.5p$.

Thus, $P[B] = P[B \cap A] + P[B \cap A'] = q + 0.5 \times p = q + 0.5 \times (1 - q) = 0.5 \times (1 + q)$,
and $P[A|B] = \frac{q}{0.5 \times (1+q)}$.
This can be summarized in the following table

| | $A =$ identical, $P[A] = q$ | $A' =$ not identical $P[A'] = 1 - q$ |
|---|---|---|
| $B =$ same sex | $P[B\|A] = 1$ (given), $P[B \cap A]$ $= P[B\|A] \times P[A] = q$ | $P[B\|A'] = 0.5$ (given) $P[B \cap A']$ $= P[B\|A'] \times P[A'] = 0.5 \times (1 - q)$ |

$$\Downarrow$$

$$P[B] = P[B \cap A] + P[B \cap A'] = q + 0.5 \times (1 - q) = 0.5 \times (1 + q).$$

Then, $P[A|B] = \frac{P[B \cap A]}{P[B]} = \frac{q}{0.5 \times (1+q)}$.

The event tree shown on page 61 can be applied to this example. □

**Bayes' rule and Bayes' Theorem (extended form):**

If $A_1, A_2, ..., A_n$ form a partition of the entire probability space $S$, then

$$P[A_j|B] = \frac{P[B\cap A_j]}{P[B]} = \frac{P[B\cap A_j]}{\sum\limits_{i=1}^{n} P[B\cap A_i]} = \frac{P[B|A_j]\times P[A_j]}{\sum\limits_{i=1}^{n} P[B|A_i]\times P[A_i]} \quad \text{for each } j = 1, 2, ..., n.$$

For example, if the $A$'s form a partition of $n = 3$ events, we have

$$P[A_1|B] = \frac{P[A_1\cap B]}{P[B]} = \frac{P[B|A_1]\times P[A_1]}{P[B\cap A_1]+P[B\cap A_2]+P[B\cap A_3]}$$
$$= \frac{P[B|A_1]\times P[A_1]}{P[B|A_1]\times P[A_1]+P[B|A_2]\times P[A_2]+P[B|A_3]\times P[A_3]}$$

The relationship in the denominator, $P[B] = \sum\limits_{i=1}^{n} P[B|A_i] \times P[A_i]$ is the general Law of Total Probability. The values of $P[A_j]$ are called prior probabilities, and the value of $P[A_j|B]$ is called a posterior probability. The basic form of Bayes' rule is just the case in which the partition consists of two events, $A$ and $A'$. The main application of Bayes' rule occurs in the situation in which the $P[A_i]$ probabilities are known and the $P[B|A_h]$ probabilities are known, and we are asked to find $P[A_j|B]$ for one of the $j$'s. The series of calculations can be summarized in a table as in the basic form of Bayes' rule. This is illustrated in the following example.

**Example 2-6:**

Three dice have the following probabilities of throwing a "six": $p, q, r$, respectively. One of the dice is chosen at random and thrown (each is equally likely to be chosen). A "six" appeared. What is the probability that the die chosen was the first one?

**Solution:**

The event " a 6 is thrown" is denoted by $B$ and $A_1$, $A_2$ and $A_3$ denote the events that die 1, die 2 and die 3 was chosen.

$$P[A_1|B] = \frac{P[A_1\cap B]}{P[B]} = \frac{P[B|A_1]\times P[A_1]}{P[B]} = \frac{p\times\frac{1}{3}}{P[B]}.$$

But $P[B] = P[B\cap A_1] + P[B\cap A_2] + P[B\cap A_3]$

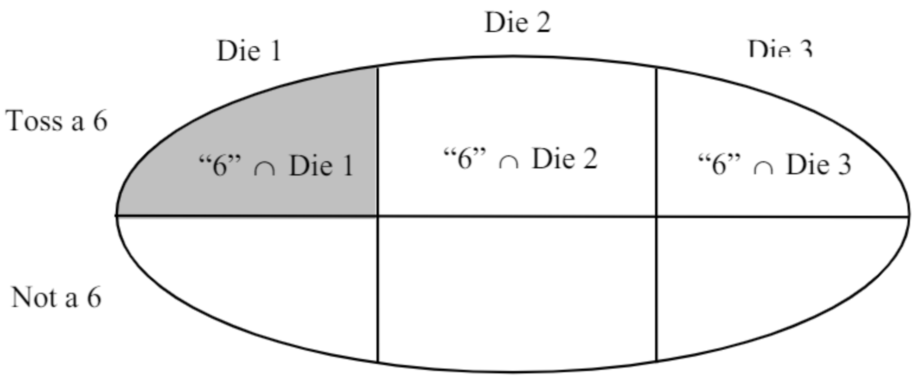
$$= P[B|A_1]\times P[A_1] + P[B|A_2]\times P[A_2] + P[B|A_3]\times P[A_3]$$
$$= p\times\tfrac{1}{3} + q\times\tfrac{1}{3} + r\times\tfrac{1}{3} = \frac{p+q+r}{3} \;\Rightarrow\; P[A_1|B] = \frac{p\times\frac{1}{3}}{P[B]} = \frac{p\times\frac{1}{3}}{(p+q+r)\times\frac{1}{3}} = \frac{p}{p+q+r}.$$

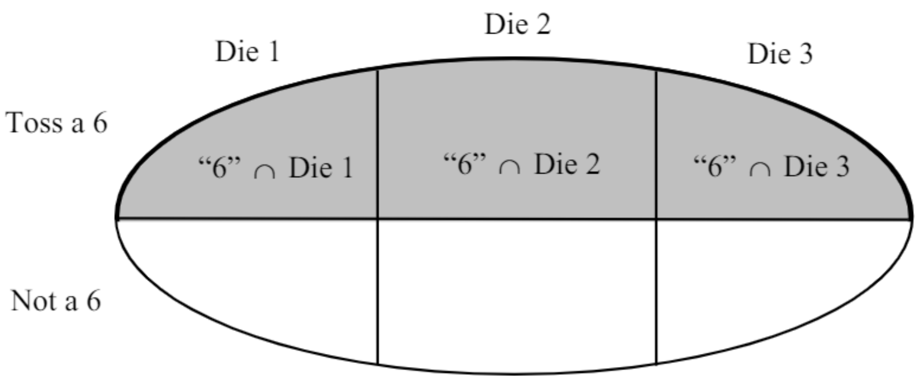
These calculations can be summarized in the following table.

| | Die 1 , $P(A_1) = \frac{1}{3}$ (given) | Die 2 , $P(A_2) = \frac{1}{3}$ (given) | Die 3 , $P(A_3) = \frac{1}{3}$ (given) |
|---|---|---|---|
| Toss "6", $B$ | $P[B\|A_1] = p$ (given)<br>$P[B\cap A_1]$<br>$= P[B\|A_1]\times P[A_1]$<br>$= p\times\frac{1}{3}$ | $P[B\|A_2] = q$ (given)<br>$P[B\cap A_2]$<br>$= P[B\|A_2]\times P[A_2]$<br>$= q\times\frac{1}{3}$ | $P[B\|A_3] = r$ (given)<br>$P[B\cap A_3]$<br>$= P[B\|A_3]\times P[A_3]$<br>$= r\times\frac{1}{3}$ |

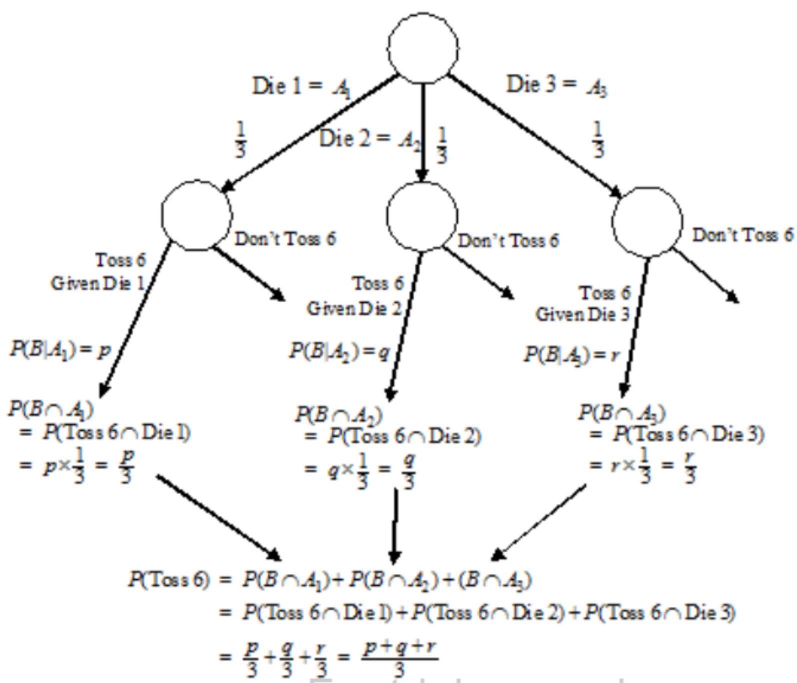
$P[B] = p\times\frac{1}{3} + q\times\frac{1}{3} + r\times\frac{1}{3} = \frac{1}{3}\times(p+q+r).$

In terms of Venn diagrams, the conditional probability is the ratio of the shaded area probability in the first diagram to the shaded area probability in the second diagram.




The event tree representing the probabilities has three branches at the top node to represent the three die types that can be chosen in the first step of the process.




□

In Example 2-6 there is a certain symmetry to the situation and general reasoning can provide a shortened solution. In the conditional probability $P[\text{die 1}|\text{"6"}] = \frac{P[(\text{die 1})\cap(\text{"6"})]}{P[\text{"6"}]}$, we can think of the denominator as the combination of the three possible ways a "6" can occur, $p+q+r$, and we can think of the numerator as the "6" occurring from die 1, with probability $p$. Then the conditional probability is the fraction $\frac{p}{p+q+r}$. The symmetry involved here is in the assumption that each die was equally likely to be chosen, so there is a $\frac{1}{3}$ chance of any one die being chosen. This factor of $\frac{1}{3}$ cancels in the numerator and denominator of $\frac{p\times\frac{1}{3}}{(p+q+r)\times\frac{1}{3}}$. If we had not had this symmetry, we would have to apply different "weights" to the three dice.

Another example of this sort of symmetry is a variation on Example 2-3 above. Suppose that Urn I has 2 white and 3 black balls and Urn II has 4 white and 1 black balls. An Urn is chosen at random and a ball is chosen. The reader should verify using the usual conditional probability rules that the probability of choosing a white is $\frac{6}{10}$. This can also be seen by noting that if we consider the 10 balls together, 6 of them are white, so that the chance of picking a white out of the 10 is $\frac{6}{10}$. This worked because of two aspects of symmetry, equal chance for picking each Urn, and same number of balls in each Urn.

**Independent events $A$ and $B$:** If events $A$ and $B$ satisfy the relationship $P[A\cap B] = P[A]\times P[B]$, then the events are said to be independent or stochastically independent or statistically independent. The independence of (non-empty) events $A$ and $B$ is equivalent to $P[A|B] = P[A]$, and also is equivalent to $P[B|A] = P[B]$.

**Example 2-1 (continued):**

A fair six-sided die is tossed.

$A =$ "the number tossed is even" $= \{2,4,6\}$ , $B =$ "the number tossed is $\leq 3$" $= \{1,2,3\}$ ,

$C =$ "the number tossed is a 1 or a 2" $= \{1,2\}$ ,

$D =$ "the number tossed doesn't start with the letters 'f' or 't'" $= \{1,6\}$ .

We have the following probabilities: $P[A] = \frac{1}{2}$, $P[B] = \frac{1}{2}$, $P[C] = \frac{1}{3}$, $P[D] = \frac{1}{3}$.

Events $A$ and $B$ are not independent since $\frac{1}{6} = P[A\cap B] \neq P[A]\times P[B] = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$. We also see that $A$ and $B$ are not independent because $P[B|A] = \frac{1}{3} \neq \frac{1}{2} = P[B]$.

Also, $B$ and $C$ are not independent, since $P[B\cap C] = \frac{1}{3} \neq \frac{1}{2}\times\frac{1}{3} = P[B]\times P[C]$ (also since $P[B|C] = 1 \neq \frac{1}{2} = P[B]$). Events $A$ and $C$ are independent, since $P[A\cap C] = \frac{1}{6} = \frac{1}{2}\times\frac{1}{3} = P[A]\times P[C]$ (alternatively, $P[A|C] = \frac{1}{2} = P[A]$, so that $A$ and $C$ are independent).

The reader should check that both $A$ and $B$ are independent of $D$. □

**Mutually independent events:** Events $A_1, A_2, \ldots, A_n$ are said to be mutually independent if the following relationships are satisfied. For any two events, say $A_i$ and $A_j$, we have $P(A_i \cap A_j) = P(A_i) \times P(A_j)$. For any three events, Say $A_i, A_j, A_k$, we have $P(A_i \cap A_j \cap A_k) = P(A_i) \times P(A_j) \times P(A_k)$. This must be true for any four events, or five events, etc.

## Some rules concerning conditional probability and independence are:

**(i)** $P[A \cap B] = P[B|A] \times P[A] = P[A|B] \times P[B]$ for any events $A$ and $B$

**(ii)** If $P[A_1 \cap A_2 \cap \cdots \cap A_{n-1}] > 0$ , then
$P[A_1 \cap A_2 \cap \cdots \cap A_n] = P[A_1] \times P[A_2|A_1] \times P[A_3|A_1 \cap A_2] \times \cdots \times P[A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}]$

**(iii)** $\boldsymbol{P[A'|B] = 1 - P[A|B]}$

**(iv)** $P[A \cup B|C] = P[A|C] + P[B|C] - P[A \cap B|C]$

**(v)** if $A \subset B$ then $P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[A]}{P[B]}$ , and $P[B|A] = 1$; properties (iv) and (v) are the same properties satisfied by unconditional events

**(vi)** if $A$ and $B$ are independent events then $A'$ and $B$ are independent events, $A$ and $B'$ are independent events, and $A'$ and $B'$ are independent events

**(vii)** since $P[\emptyset] = P[\emptyset \cap A] = 0 = P[\emptyset] \times P[A]$ for any event $A$, it follows that $\emptyset$ is independent of any event $A$

**Example 2-7:**

Suppose that events $A$ and $B$ are independent. Find the probability, in terms of $P[A]$ and $P[B]$, that exactly one of the events $A$ and $B$ occurs.

**Solution:** $P[\text{exactly one of } A \text{ and } B] = P[(A \cap B') \cup (A' \cap B)]$.
Since $A \cap B'$ and $B \cap A'$ are mutually exclusive, it follows that
$P[\text{exactly one of } A \text{ and } B] = P[A \cap B'] + P[A' \cap B]$.
Since $A$ and $B$ are independent, it follows that $A$ and $B'$ are also independent, as are $B$ and $A'$.
Then $P[(A \cap B') \cup (A' \cap B)] = P[A] \times P[B'] + P[B] \times P[A']$
$= P[A] \times (1 - P[B]) + P[B] \times (1 - P[A]) = P[A] + P[B] - 2P[A] \times P[B]$ □

**Example 2-8:**

In a survey of 94 students, the following data was obtained.

60 took English, 56 took Math, 42 took Chemistry, 34 took English and Math, 20 took Math and Chemistry, 16 took English and Chemistry, 6 took all three subjects.

Find the following proportions.

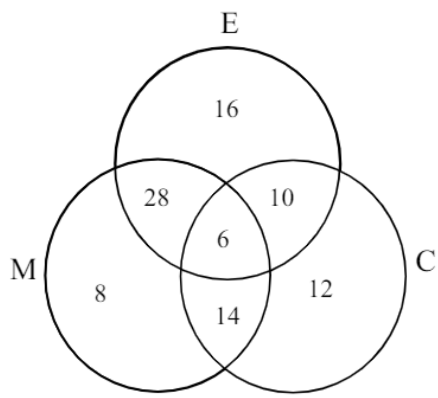
(i) Of those who took Math, the proportion who took neither English nor Chemistry,

(ii) Of those who took English or Math, the proportion who also took Chemistry.

**Solution:**

The following diagram illustrates how the numbers of students can be deconstructed.

We calculate proportion of the numbers in the various subsets.




(i) 56 students took Math, and 8 of them took neither English nor Chemistry.
$P(E' \cap C'|M) = \frac{P(E' \cap C' \cap M)}{P(M)} = \frac{8}{56} = \frac{1}{7}$.

(ii) 82 $(= 8 + 14 + 6 + 28 + 16 + 10$ in $E \cup M)$ students took English or Math (or both), and 30 of them $(= 14 + 6 + 10$ in $(E \cup M) \cap C)$ also took Chemistry.
$P(C|E \cup M) = \frac{P[C \cap (E \cup M)]}{P(E \cup M)} = \frac{30}{82} = \frac{15}{41}$. □

**Example 2-9:**

A survey is made to determine the number of households having electric appliances in a certain city. It is found that 75% have radios $(R)$, 65% have irons $(I)$, 55% have electric toasters $(T)$, 50% have $(IR)$, 40% have $(RT)$, 30% have $(IT)$, and 20% have all three.

Find the following proportions.

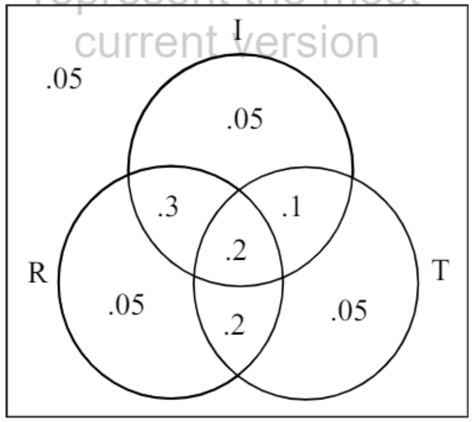
(i) Of those households that have a toaster, find the proportion that also have a radio.

(ii) Of those households that have a toaster but no iron, find the proportion that have a radio.

**Solution:**

This is a continuation of Example 1-3 given earlier in the study guide.

The diagram below deconstructs the three events.




(i) This is $P[R|T]$. The language "of those households that have a toaster" means, "given that the household has a toaster", so we are being asked for a conditional probability.
Then, $P[R|T] = \frac{P[R \cap T]}{P[T]} = \frac{0.4}{0.55} = \frac{8}{11}$.

(ii) This is $P[R|T \cap I'] = \frac{P[R \cap T \cap I']}{P[T \cap I']} = \frac{.2}{.25} = \frac{4}{5}$. □

Example 2-8 presents a "population" of 94 individuals, each with some combination of various properties (took English, took Math, took Chemistry). We found conditional probabilities involving the various properties by calculating proportions in the following way

$$P(A|B) = \frac{number\,of\,individuals\,satisfying\,both\,properties\,A\,and\,B}{number\,of\,indiviudals\,satisfying\,property\,B}$$

We could approach Example 2-9 in a similar way by creating a "model population" with the appropriate attributes. Since we are given percentages of households with various properties, we can imagine a model population of 100 households, in which 75 have radios ($R$, 75%), 65 have irons ($I$), 55 have electric toasters ($T$), 50 have ($IR$), 40 have ($RT$), 30 have ($IT$), and 20 have all three. The diagram in the solution could be modified by changing the decimals to numbers out of 100, so .2 becomes 20, etc. Then so solve (i), since 55 have toasters and 40 have both a radio and a toaster, the proportion of those who have toasters that also have a radio is $\frac{40}{55}$.

Creating a model population is sometimes an efficient way of solving a problem involving conditional probabilities, particularly when applying Bayes rule. The following example illustrates this.

**Example 2-10 (SOA):**

A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease. Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

**Solution:**

We identify the following events:

$D$ : a person has the disease ,

$TP$ : a person tests positive for the disease

We are given

$P(D) = 0.01,$

$P(D') = 0.99,$

$P(TP|D) = 0.95,$

$P(TP'|D) = 0.05,$

$P(TP|D') = 0.005$

$P(TP'|D') = 0.995$ .

We wish to find $P(D|TP)$. We first solve the problem using rules of conditional probability.
We have $P(D|TP) = \frac{P(D \cap TP)}{P(TP)}$.
We also have, $P(D \cap TP) = P(TP|D) \times P(D) = 0.95 \times 0.01 = 0.0095$, and

$P(TP) = P(D \cap TP) + P(D' \cap TP)$

$= P(TP|D) \times P(D) + P(TP|D') \times P(D') = 0.95 \times 0.01 + 0.005 \times .99 = \times .01445.$

Then, $P(D|TP) = \frac{P(D \cap TP)}{P(TP)} = \frac{0.0095}{0.01445} = 0.657.$

We can also solve this problem with the model population approach. We imagine a model population of 100,000 individuals. In this population, the number with disease is $\#(D) = 1000$ (.01 of the population) , the number without disease is $\#(D') = 99,000$ (0.99 of the population).

Since $P(TP|D) = 0.95$ , it follows that 95% of those with the disease will test positive, so the number who have the disease and test positive is $\#(TP \cap D) = 0.95 \times 1000 = 950$ (this just reflects the fact that $P(TP \cap D) = P(TP|D) \times P(D) = 0.95 \times 0.01 = .0095$ , so that $0.0095 \times 100,000 = 950$ in the population have the disease and test positive. In the same way, we find $\#(TP \cap D') = 0.005 \times 99,000 = 495$ is the number who do not have disease but test positive. Therefore, the total number who test positive is

$$\#(TP) = \#(TP \cap D) + \#(TP \cap D') = 950 + 495 = 1445.$$

The probability that a person has the disease given that the test indicates the presence of the disease is the proportion that have the disease and test positive as a fraction of all those who test positive,

$$P(D|TP) = \frac{\#(TP \cap D)}{\#(TP)} = \frac{950}{1445} = 0.657.$$

The following table summarizes the calculations in the conditional probability approach.

| | | |
|---|---|---|
| | $P(D) = 0.01$ , given | $P(D') = 0.99$ <br> $= 1 - 0.01$ |
| $TP$ | $P(TP\|D) = 0.95$, given <br> $P(TP \cap D)$ <br> $= P(TP\|D) \times P(D)$ <br> $= 0.95 \times 0.01 = 0.0095$ | $P(TP\|D') = 0.005$ , given <br> $P(TP \cap D')$ <br> $= P(TP\|D') \times P(D')$ <br> $= 0.005 \times 0.99 = 0.00495$ |
| $TP'$ | $P(TP'\|D) = 1 - P(TP\|D)$ <br> $= 0.05$ , <br> $P(TP' \cap D)$ <br> $= P(TP'\|D) \times P(D)$ <br> $= 0.05 \times 0.01 = 0.0005$ | $P(TP'\|D') = 1 - P(TP\|D')$ <br> $= 0.995$ , <br> $P(TP' \cap D')$ <br> $= P(TP'\|D') \times P(D')$ <br> $= 0.995 \times 0.99 = 0.98505$ |

$P(TP) = P(TP \cap D) + P(TP \cap D') = 0.0095 + 0.00495 = 0.01445.$

$P(D|TP) = \frac{P(D \cap TP)}{P(TP)} = \frac{0.0095}{0.01445} = 0.657.$ $\square$