# Chapter 5

## Discussion Question Solutions

**D1.** Necker's model is based on a list of four possible outcomes. There is exactly one outcome with 0 heads, two outcomes with 1 head, and one outcome with 2 heads. Hence, the respective probabilities are $\frac{1}{4}$, $\frac{2}{4}=\frac{1}{2}$, and $\frac{1}{4}$.

**D2.** The sum of each set of probabilities is 1. The sum of the probabilities must be 1 because the distribution includes all possible disjoint outcomes for a random process and one of them must occur.

**D3.** Necker is correct; d'Alembert did not account for the two ways that a first flip of tails, namely *TH* and *TT*. To decide who is right, your class could simulate the situation by flipping two distinct coins, perhaps a penny and a nickel. Flip the two coins many times and count the outcomes, {0 heads, 1 head, 2 heads}. Decide whether the proportions of times the outcomes occur are closer to d'Alembert's or closer to Necker's.

**D4.** Following the example in the text and Display 5.3 with two testers, a three-tester sample space would look like

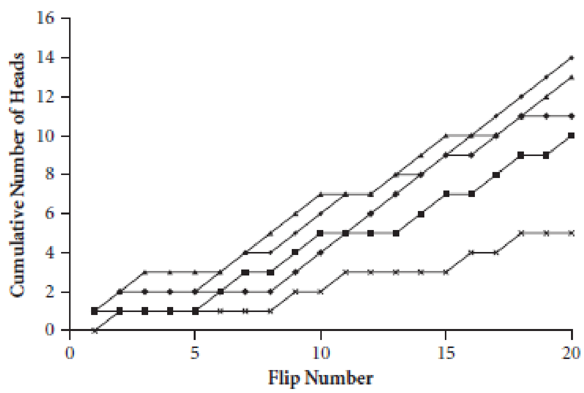
| Taste Tester A | Taste Tester B | Taste Tester C |
|---|---|---|
| G | G | G |
| G | G | O |
| G | O | G |
| G | O | O |
| O | G | G |
| O | G | O |
| O | O | G |
| O | O | O |

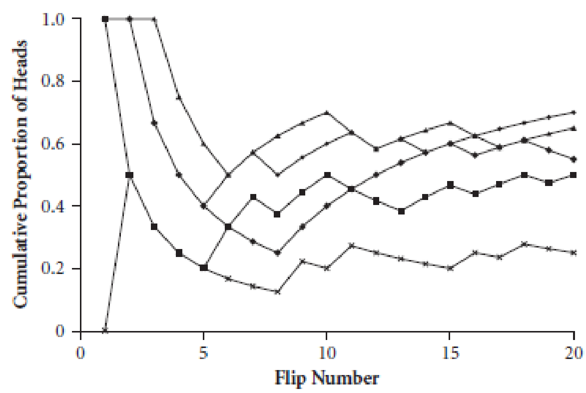
where G indicates the tester identified the actual gourmet coffee correctly and O indicates the tester identified the ordinary coffee incorrectly as the gourmet coffee. These outcomes will be equally likely if all three testers are truly guessing. That is, each tester is just as likely to be right as wrong. All three testers selecting the correct coffee should not be enough to convince a skeptic that people can distinguish between types of coffee. If all three tasters were just guessing, the outcome of all three being correct has a probability of 1/8 which is fairly large.

**D5.** If Taste Tester A truly has the ability to identify the gourmet coffee most of the time, then the outcomes will not be equally likely. Any outcome in the first four rows of the table in D4 above, with G in the first column, will be more likely than those in the last four rows.

**D6.** The pollster is correct—but only if the sampling is done randomly. Then the Law of Large Numbers practically guarantees that in large samples the sample proportions will be close to the corresponding population proportions. The casino operator is correct because the machines and games have a chance mechanism built in, and the probability of the house winning is always greater than 0.5. If many people play, the Law of Large Numbers takes over and the house is practically guaranteed to make money. The manufacturer is wrong. Increasing the volume will increase the number of good products produced, as well as the number of defectives, but increasing the volume alone will not change the workmanship by rushed workers. With machinery, the fraction of defectives should converge to a constant that represents the machine's true probability of producing a defective product.

**D7.** The plots below show the cumulative counts and the cumulative proportions for five sets of 20 flips each. As the number of flips increases, the counts of heads diverge while the proportions of heads converge. (Convergence is not strongly apparent, however, for samples of only 20.) The fact that sample proportions (and sample means) converge to a constant makes them quite useful as estimators of population quantities.




**D8. a**. There are $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$, or 128, possible outcomes.

**b.** There is only one way to get seven heads, so the probability is $\frac{1}{128}$.

**c.** There are exactly seven ways to get six heads and one tail (the tail could occur on any one of the seven flips), so the probability is $\frac{7}{128}$.

**D9.** There are $3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 243$ possible outcomes because each of the five people could make any one of three choices (Starbucks, McDonald's, Dunkin' Donuts). The equally likely model is not a good one in this case because there is no reason to believe that the three choices are equally likely.

**D10. a.** Disjoint. If a person is not in the labor force, then they are neither employed nor seeking employment so the last category is certainly disjoint from the others. Further if someone is unemployed then they cannot be in either of the first two categories of employed people. The only possibility for overlap is in the first two categories as someone with two jobs could be employed in both the agricultural and nonagricultural industries. However, we are told only the principal source of employment was counted so no respondent would be in both categories.

**b.** Not disjoint. Nothing about the categories themselves indicates disjointedness since, for example, a person who dines out may certainly read books or play computer games or go to the beach. Proof that the categories are not disjoint can be seen by simply adding the percentages. Since $49\% + 39\% + 20\% + 24\% = 132\% > 100\%$, there must be respondents in multiple categories.

**D11.** Only the events in parts c and e must be mutually exclusive.

**D12.** No, although

$P(\textit{get bitten}) = P(\textit{get bitten the first time you go outside}$ or $\textit{get bitten the second time you go outside}$ or $\textit{get bitten the third time you go outside})$,

it is not true that this is equal to

$P(\textit{get bitten the first time you go outside}) + P(\textit{get bitten the second time you go outside}) + P(\textit{get bitten the third time you go outside})$.

That's because the three events are not disjoint. You could get bitten each time you went outside. If you go outside six times, it is clear that the probability you will get bitten cannot be $0.2+0.2+0.2+0.2+0.2+0.2=1.2>1$.

**D13. a.** Diagram II illustrates disjoint events. The regions share no area, meaning both events cannot happen on the same occurrence.

**b.** Ann and Bill incorrectly thought that diagram II applied, that is, that they had no close friends in common. Diagram I actually applies, they do have nearest and dearest friends in common. If they had no friends in common, the number of people at the party would be 200. Since only 140 people were there, they must have $200-140=60$ friends in common and each has $100-60=40$ friends that are not friends of the other.

**c.** A Venn diagram for this situation would look like diagram I where *A* represents those who favor state recognition and *B* those who favor federal recognition. There may be those who favor one and not the other as well as those that favor both. Without further information, such as the percent of those that favor both or the percent that favor at least one, there is no way to determine the percent in each section of the diagram.

**D14.** If *A* and *B* are mutually exclusive, then $P(A \text{ and } B) = 0$, so the general form of the Addition Rule reduces to the Addition Rule for Mutually Exclusive Events.

**D15. a.** The conditional probability is $\frac{1}{3}$ without replacement and $\frac{1}{2}$ with replacement. These values are quite far apart. b. The conditional probability is $\frac{49}{99}$ without replacement and $\frac{1}{2}$ with replacement. These values are quite close. c. If the population size is large relative to the sample size, sampling without replacement is about the same as sampling with replacement. If the population is small relative to the sample size, sampling without replacement can change the probabilities significantly.

**D16.**

$$P(M \text{ and } S) = P(M)\cdot P(S \mid M) = \frac{1731}{2201}\cdot\frac{367}{1731} = \frac{367}{2201}$$

$$P(M \text{ and } D) = P(M)\cdot P(D \mid M) = \frac{1731}{2201}\cdot\frac{1364}{1731} = \frac{1364}{2201}$$

**D17.**

$$P(M \text{ and } S) = \frac{367}{2201} = \frac{711}{2201}\cdot\frac{367}{711} = P(S)\cdot P(M \mid S)$$

$$P(M \text{ and } S) = \frac{367}{2201} = \frac{1731}{2201}\cdot\frac{367}{1731} = P(M)\cdot P(S \mid M)$$

**D18.** Let $n = c + d + e + f$. Then, $P(A)\cdot P(B \mid A) = \frac{(c+e)}{n}\cdot\frac{c}{(c+e)} = \frac{c}{n} = P(A \text{ and } B)$

and $P(B)\cdot P(A \mid B) = \frac{(c+d)}{n}\cdot\frac{c}{(c+d)} = \frac{c}{n} = P(A \text{ and } B)$

**D19.** A person who wants to interpret test results usually has been told by a doctor that he or she has a positive test result or a negative test result. They want to know the probability that the test result is correct. If, for example, someone has a positive test result, the PPV gives them some idea of the probability that they really have the disease. On the other hand, sensitivity is the probability the test will be positive for a randomly selected person who has the disease. For an individual wondering whether he or she has the disease, knowing the sensitivity is not very helpful. As students saw in the example A Relatively Rare Disease on page 252, the PPV, the proportion of people who test positive who actually have the disease, can be fairly low when the incidence of the disease in the population is low. High numbers of false positives are one reason that results of screening tests for some serious conditions, such as HIV, are given only in person.

**D20.** PPV: $P(\textit{actually guilty} \mid \textit{found guilty})$ NPV: $P(\textit{actually not guilty} \mid \textit{found not guilty})$
The legal system is set up to keep the number of false positives to a minimum and so the PPV to a maximum.

**D21. a.** Consider the two events *Tester A selects O* and *Tester B selects O*. Then from the sample space

$$P(\textit{Tester B selects O}) = \frac{1}{2}$$

and

$$P(Tester\ B\ selects\ O \mid Tester\ A\ selects\ O) = \frac{P(Tester\ B\ selects\ O \text{ and } Tester\ A\ selects\ O)}{P(Tester\ A\ selects\ O)} = \frac{1/4}{2/4} =$$

The probabilities are equal and so the two events are independent.

**b.** Consider the two events *A selected vice-chair* and *C selected chair*. Then from the sample space

$$P(A\ selected\ vice\text{-}chair) = \frac{2}{6} = \frac{1}{3}$$

and

$$P(A\ selected\ vice\text{-}chair \mid C\ selected\ chair) = \frac{P(A\ selected\ vice\text{-}chair \text{ and } C\ selected\ chair)}{P(C\ selected\ chair)} = \frac{1}{2}$$

The probabilities are *not* equal and so the two events are independent.

**D22.** Note that you are comparing $P(B \mid A)$ and $P(B)$.
**a.** *A* happening should decrease the probability of *B*.
**b.** *A* happening should increase the probability of *B*.
**c.** *A* happening should not change the probability of *B*.
**d.** *A* happening decreases the probability of *B* to zero.

**D23. a.** This situation is not possible. Suppose that events *A* and *B* are disjoint. Then $P(A \mid B) = 0.$ If *A* and *B* are also independent, this implies that $P(A) = 0.$ But this contradicts the definition of independence on page 259 in the student book, which states that $P(A) \neq 0.$

**b.** This situation is possible. For example, when rolling two dice, a sum of 7 and a 1 on the first die are not disjoint events and are independent.

**c.** This is possible. For example, rolling a sum of 7 and rolling doubles are disjoint, dependent events.

**d.** This is possible. For example, rolling doubles and rolling a sum of 8 are not disjoint events and are dependent.

**D24.** $P(all\ three\ below\ median\ price) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

$$P(at\ least\ one\ below\ median\ price) = 1 - P(all\ three\ below\ median\ price)$$
$$= 1 - \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{7}{8}$$

This solution assumes that a large number of houses were sold recently so that sampling without replacement can be modeled as sampling with replacement.

**D25. a.** The probability of a crib death for a family such as the Clarks was estimated to be $\frac{1}{8500}$. To estimate the probability of two crib deaths in such a family, the pediatrician multiplied this probability by itself

$$\frac{1}{8500}\cdot\frac{1}{8500}=\frac{1}{72,250,000}$$

or approximately 1 in 73 million. In order to multiply the probabilities, the pediatrician assumed that a second infant death in a household is independent of a first death in the same household. Actual data did not bear out this assumption, a second death is much more likely given a first death.

**b.** The death records supplied an estimate of the probability of a two-death family as

$$\frac{1}{1303}\cdot\frac{1}{100}=\frac{1}{130,300}$$

while the number of second births per year is 220,000. The expected number of two-death families is then

$$\frac{1}{130,300}\cdot 220,000\approx 1.69$$

so between one and two two-death families will be seen each year on average.

**c.** Although a probability of 1 in 73 million is small, if one considers all second births in the world and not just in England, the expected number per year would be larger. The Clark family's medical history should be considered to see if they were more likely to experience a crib death. Lastly, this probability should not have been the most damaging piece of evidence. Motive and means would need to be considered.

# Practice Problem Solutions

**P1. a.** Out of the 27 days listed on which the National Weather Service forecast a low temperature of 30$^{o}$F, 13 of them actually had a low temperature of approximately 30$^{o}$. Based on this, the best estimate of the probability would be 13/27, or about 0.48. Also, for 25$^{o}$F, the best estimate of the probability would be 8/27, or about 0.30.

**b.** There appears to be a "cold" bias. When the prediction was incorrect the actual temperature was colder than predicted 10 times compared to only four times the temperature was warmer than predicted.

**P2. a.** There are 32 possible outcomes:

| 5 Heads | 4 Heads | 3 Heads | 2 Heads | 1 Head | 0 Heads |
|---|---|---|---|---|---|
| *HHHHH* | *HHHHT* | *HHHTT* | *HHTTT* | *HTTTT* | *TTTTT* |
| | *HHHTH* | *HHTHT* | *HTHTT* | *THTTT* | |
| | *HHTHH* | *HTHHT* | *THHTT* | *TTHTT* | |
| | *HTHHH* | *THHHT* | *HTTHT* | *TTTHT* | |
| | *THHHH* | *HHTTH* | *THTHT* | *TTTTH* | |
| | | *HTHTH* | *TTHHT* | | |
| | | *THHTH* | *HTTTH* | | |
| | | *HTTHH* | *THTTH* | | |
| | | *THTHH* | *TTHTH* | | |
| | | *TTHHH* | *TTTHH* | | |

**b.** Yes, all outcomes are equally likely.

**c.** For these, simply count the number of outcomes satisfying each condition and divide by 32. In the order presented, these probabilities are: 1/32, 5/32, 10/32, 10/32, 5/32, 1/32

**P3. a.** One way to do this is shown here:

H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6

**b.** Answers will vary. For the sample space shown in part A the outcomes are equally likely.

**c.** 1/12

**P4. a.**

| | | |
|---|---|---|
| 28, 35 | 28, 39 | 28, 47 |
| 28, 55 | 35, 39 | 35, 47 |
| 35, 55 | 39, 47 | 39, 55 |
| 47, 55 | | |

**b.** Yes.

**c.** 1/10

**d.** 6/10

**P5. a.** Label each person using the first letter of their first name. The sample space consists of the outcomes AB, AC, AD, BC, BD, CD.

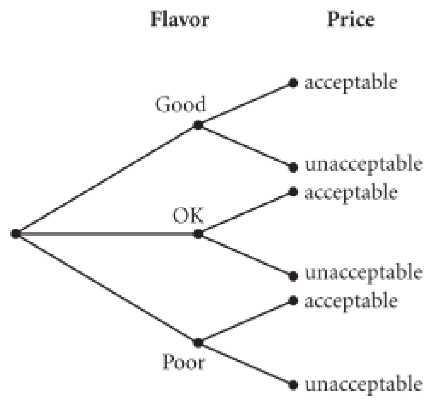
**b.** Count all outcomes that contain an A, and divide by 6 to conclude that the probability is 3/6, or ½.

**P6. a.** Yes, you can list a sample space and it would look similar to the 32 outcomes in P1, with *T* representing being right-handed and *B* representing being left-handed.

**b.** You can not determine the probability without additional information about the percentage of students in the school who are right-handed.

**P7. a.** Since the proportion that was heads after the first two spins was zero, the first and second spins must have been tails. The third was heads because the proportion that were heads went up. The fourth was tails and the $50^{th}$ was tails.
**b.** About 0.44.

**P8. a.**




**b.** Six.
**c.** It is impossible to tell the probabilities without knowing the quality of the ice cream and price, but it is unlikely that the outcomes are equally likely.

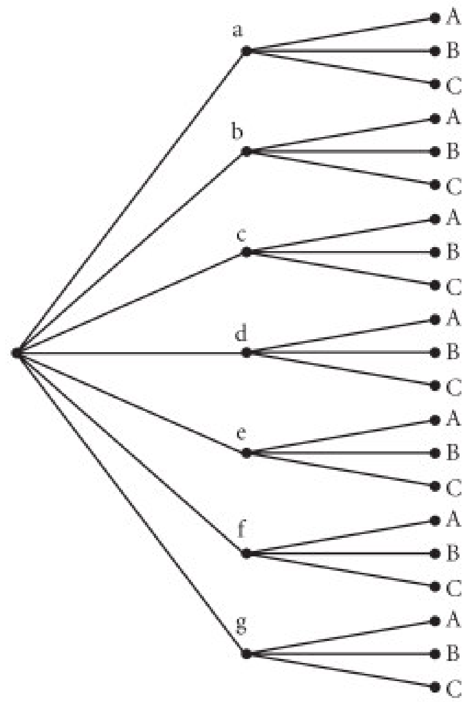
**P9. a.** There are 3 · 7, or 21, different pairs of "levels of exercise" and "different medications" that you could end up with.

**b.** There is only one element of the sample space that corresponds to this event, so its probability is $\frac{1}{21}$.

**c.** Label the medications as "a, b, c, d, e, f, g" and the levels of exercises as "A, B, C". The possible outcomes, in array form, are:

| | A | B | C |
|---|---|---|---|
| a | aA | aB | aC |
| b | bA | bB | bC |
| c | cA | cB | cC |
| d | dA | dB | dC |
| e | eA | eB | eC |
| f | fA | fB | fC |
| g | gA | gB | gC |

**d.** Using the same labeling scheme as in (c), the possible outcomes, in tree form, are:




**P10. a.** $\frac{32.3+4.9}{298.6} \approx 0.125$

**b.** i. These are not disjoint since there are females 85 or older in the sample.

ii. These are disjoint since one cannot be $\leq 17$ and $\geq 85$ years of age simultaneously.

iii. These are not disjoint since a person in the "$\geq 85$ years of age" category is also in the "$\geq 65$ years of age" category.

**P11. a.** Use the information in the rightmost column since gender is not specified:

$$\frac{6841+3958}{17,231} \approx 0.627$$

**b.** Since the gender specified is male, use the leftmost column:

$$\frac{3171+1782}{7505} \approx 0.660$$

**c.** Since the gender specified is female, use the rightmost column:

$$\frac{3670+2176}{9726} \approx 0.60$$

**P12. a.** No **b.** Yes **c.** Yes **d.** No

**P13. a.** Yes, because selecting a junior and selecting a senior are disjoint events.

**b.** 33% + 27% = 60%

**P14. a.**

| | | Second Roll | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| First Roll | 1 | 1, 1 | 1, 2 | 1, 3 | 1, 4 |
| | 2 | 2, 1 | 2, 2 | 2, 3 | 2, 4 |
| | 3 | 3, 1 | 3, 2 | 3, 3 | 3, 4 |
| | 4 | 4, 1 | 4, 2 | 4, 3 | 4, 4 |

**b.** P(6 or 7) = P(6) + P(7) = 3/16 + 2/16 = 5/16.
**c.** P(doubles or 7) = P(doubles) + P(7) = 4/16 + 2/16 = 6/16
**d.** Because these events are not disjoint. The outcome 3,3 is both a double and a 6.

**P15. a.** No. If they were mutually exclusive the cell in the column for alcohol involved and the row for speed related would be 0.

**b.** $\frac{10{,}928}{231{,}459} + \frac{87{,}086}{231{,}459} - \frac{4{,}436}{231{,}459} \approx 0.404$

**P16. a.**

$$P(\textit{doubles}\text{ or }\textit{sum of }4) = P(\textit{doubles}) + P(\textit{sum of }4) - P(\textit{doubles}\text{ and }\textit{sum of }4)$$
$$= \tfrac{6}{36} + \tfrac{3}{36} - \tfrac{1}{36} = \tfrac{8}{36}$$

**b.**

$$P(\textit{doubles}\text{ or }\textit{sum of }7) = P(\textit{doubles}) + P(\textit{sum of }7) - P(\textit{doubles}\text{ and }\textit{sum of }7)$$
$$= \tfrac{6}{36} + \tfrac{6}{36} - \tfrac{0}{36} = \tfrac{12}{36}$$

**c.**

$$P(5\textit{ on first die}\text{ or }5\textit{ on second die}) = P(5\textit{ on first die}) + P(5\textit{ on second die})$$
$$- P(5\textit{ on first die}\text{ and }5\textit{ on second die})$$
$$= \tfrac{6}{36} + \tfrac{6}{36} - \tfrac{1}{36} = \tfrac{11}{36}$$

**P17.** $P(2\textit{ heads}) = P(\textit{heads on first}) + P(\textit{heads on second}) - P(\textit{heads on both})$
$= \tfrac{1}{2} + \tfrac{1}{2} - \tfrac{1}{4} = \tfrac{3}{4}$

**P18.** *P(not doubles* or *sum of 8) = P(not doubles) + P(sum of 8)*
– *P(not doubles* and *sum of 8)*
$= \tfrac{30}{36} + \tfrac{5}{36} - \tfrac{4}{36} = \tfrac{31}{36}$

**P19. a.** $P(F) = P(\textit{the person was female}) = \frac{470}{2201}$
**b.** $P(F \mid S) = P(\textit{the person was female}\text{ given that }\textit{the person survived}) = \frac{344}{711}$
**c.** $P(\textit{not } F) = P(\textit{the person wasn't female}) = \frac{1731}{2201}$
**d.** $P(\textit{not } F \mid S) = P(\textit{the person wasn't female}\text{ given that }\textit{the person survived}) = \frac{367}{711}$
**e.** $P(S \mid \textit{not } F) = P(\textit{the person survived}\text{ given that }\textit{the person was not female}) = \frac{367}{1731}$

**P20. a.** $P(\textit{worker is paid at or below minimum wage}) = \frac{1675}{73,756} \approx 0.023$

**b.** $P(\textit{worker is paid at or below minimum wage} \mid \textit{worker is white}) = \frac{1420}{61,061} \approx 0.023$

**c.** White workers have approximately the same likelihood of being paid at or below minimum wage as the population in general.

**d.** Find $P(\textit{worker is black}) = \frac{9965}{73,756} \approx 0.135$

**e.** Find $P(\textit{worker is black} \mid \textit{worker is paid at or below minimum wage}) = \frac{205}{1675} \approx 0.122$

**f.** A worker that is paid at or below minimum wage is less likely to be black than a worker selected at random from all hourly workers.

---

**P21. a.** $P(\textit{2nd draw is red} \mid \textit{1st draw is red}) = \frac{2}{4} = \frac{1}{2}$

**b.** $P(\textit{2nd draw is red} \mid \textit{1st draw is blue}) = \frac{3}{4}$

**c.** $P(\textit{3rd is blue} \mid \textit{1st is red} \text{ and } \textit{2nd is blue}) = \frac{1}{3}$

**d.** $P(\textit{3rd is red} \mid \textit{1st is red} \text{ and } \textit{2nd is red}) = \frac{1}{3}$

---

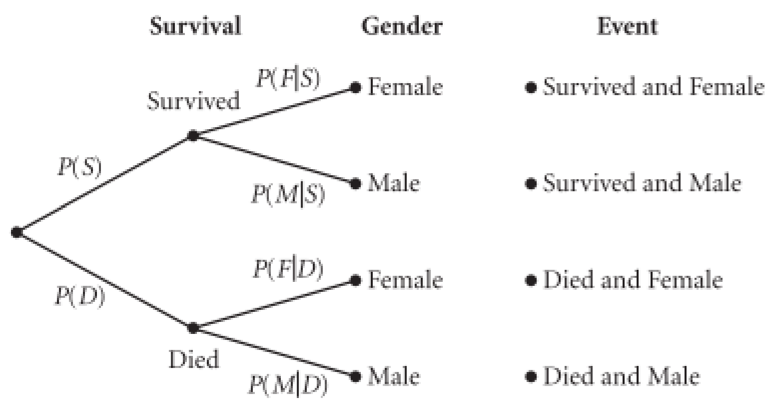
**P22. a.** $P(club \mid black) = \frac{\textit{number of black clubs}}{\textit{total number of black cards}} = \frac{13}{26} = \frac{1}{2}$

**b.** $P(jack \mid heart) = \frac{\textit{number of jack of hearts}}{\textit{total number of black cards}} = \frac{1}{13}$

**c.** $P(heart \mid jack) = \frac{\textit{number of jacks of hearts}}{\textit{total number of jacks}} = \frac{1}{4}$

---

**P23. a.**




**b.** i. $P(S) = \frac{711}{2201}$ ii. $P(F \mid S) = \frac{344}{711}$ iii. $P(S \text{ and } F) = \frac{344}{2201}$

**c.** $P(S) \cdot P(F \mid S) = P(S \text{ and } F)$. This formula matches the computation in the text.

**d.** $P(M \text{ and } S) = P(M) \cdot P(S \mid M)$ and $P(M \text{ and } S) = P(S) \cdot P(M \mid S)$

---

**P24.** Without replacement, $P(HH) = \frac{13}{52} \cdot \frac{12}{51} = \frac{1}{17}$

With replacement, $P(HH) = \frac{13}{52} \cdot \frac{13}{52} = \frac{1}{16}$

---

**P25.** $P(\textit{W chosen 1st}) = \frac{2}{4}$

$P(\textit{W chosen 2nd} \mid \textit{W chosen 1st}) = \frac{1}{3}$

$P(WW) = P(\textit{W chosen 1st}) \cdot P(\textit{W chosen 2nd} \mid \textit{W chosen 1st}) = \frac{2}{4} \cdot \frac{1}{3} = \frac{1}{6}$

**P26.** $P(\textit{sum of eight} \text{ and } \textit{doubles}) = P(\textit{doubles}) \cdot P(\textit{sum of eight} \mid \textit{doubles}) = \frac{6}{36} \cdot \frac{1}{6} = \frac{1}{36}$

Note that this is not $P(\textit{sum of eight}) \cdot P(\textit{doubles})$, which is $\frac{5}{36} \cdot \frac{6}{36} = \frac{5}{216}$.

**P27.** $\text{P}(\textit{doubles} \mid \textit{sum of 8}) = \text{P}(\textit{doubles} \text{ and } \textit{sum of 8}) / \text{P}(\textit{sum of 8}) = \frac{\frac{1}{36}}{\frac{5}{36}} = \frac{1}{5}$.

$\text{P}(\textit{sum of 8} \mid \textit{doubles}) = \text{P}(\textit{doubles} \text{ and } \textit{sum of 8}) / \text{P}(\textit{doubles}) = \frac{\frac{1}{36}}{\frac{6}{36}} = \frac{1}{6}$.

These probabilities are not the same.

**P28.**

| | Right-Handed | Left-Handed | Total |
|---|---|---|---|
| **Blue Eyes** | 8 | 2 | **10** |
| **Brown Eyes** | 16 | 4 | **20** |
| **Total** | **24** | **6** | **30** |

$$P(\textit{right}-\textit{handed} \mid \textit{brown eyes}) = \frac{P(\textit{right}-\textit{handed} \text{ and } \textit{brown eyes})}{P(\textit{brown eyes})} = \frac{16/30}{20/30} = \frac{16}{20}$$

**P29.** You can interpret the two percentages given by Gallup as the approximate conditional probabilities that a randomly selected Republican woman, *W*, or Republican man, *M*, respectively, would have voted for Dole, *D*. That is,

$$\text{P}(D \mid W) = 0.16 \quad \text{and} \quad \text{P}(D \mid M) = 0.07$$

You can write the probability of a Republican voting for Dole, P(*D*) in terms of two disjoint events as follows:

$$\begin{aligned} \text{P}(D) &= \text{P}[(D \text{ and } W) \text{ or } (D \text{ and } M)] \\ &= \text{P}(D \text{ and } W) + \text{P}(D \text{ and } M) \\ &= \text{P}(W) \cdot \text{P}(D \mid W) + \text{P}(M) \cdot \text{P}(D \mid M) \\ &= 0.40 \cdot 0.16 + 0.60 \cdot 0.07 = 0.106 \end{aligned}$$

**P30. a.**

| | | Test Result | | |
|---|---|---|---|---|
| | | Positive | Negative | Total |
| **Disease?** | Yes | 2,985 | 0.005 • 3,000 = 15 | **3,000** |
| | No | 0.06 • 97,000 = 5,820 | 91,180 | **97,000** |
| | Total | **8,805** | **91,195** | **100,000** |

**b.** $\frac{8{,}805}{100{,}000} = 0.088$

**c.** $\frac{0.0582}{0.088} \approx 0.66$ (Note that this result is the reason why many people do not believe that there should be universal, mandatory testing for the HIV virus, prostate cancer, and other such diseases. When the incidence of the disease is relatively low in the population, the number of false positives can be much larger than the number of cases of the disease.)

**P31. a.**

PPV = P( Fluid present | ultrasound positive) = $\frac{14}{29} \approx 0.483$

NPV = P(No fluid present | ultrasound negative) = $\frac{290}{299} \approx 0.970$

Sensitivity = P(ultrasound positive | fluid present) = $\frac{14}{23} \approx 0.609$

Specificity = P(ultrasound negative | fluid not present) = $\frac{290}{305} \approx 0.951$

**b.** No. The sensitivity is only 0.609. The test fails to find fluid is almost 40% of the women who have fluid. Even if positive, the more accurate test will be needed anyway because about half of the positives are false positives.

**P32. a.** $\frac{1}{2}$

**b.** 1

**c.** We cannot tell because the probability depends on the probability model used for selecting the coin; either coin could have produced the observed head.

**d.** Same as c.

**e.** 1

**f.** 0

**P33.** $P(\text{didn't survive}) = \frac{1490}{2201} \approx .677$ is not equal to $P(\text{didn't survive} \mid \text{male}) = \frac{1364}{1731} \approx .788$

So, the events *didn't survive* and *male* aren't independent. In fact, no two events described in this table are independent.

**P34.** Parts a and c give pairs of independent events, but part b does not.

**a.** $P(heart \mid jack) = \frac{1}{4} = \frac{13}{52} = P(heart)$

**b.** $P(heart \mid red\ card) = \frac{1}{2} \neq \frac{1}{4} = P(heart)$

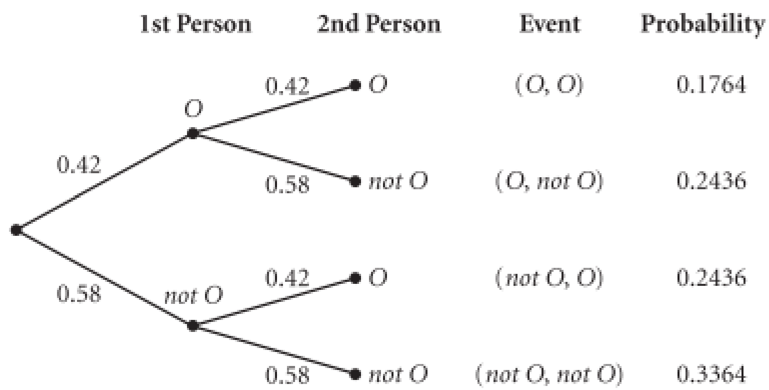
**c.** $P(getting\ a\ 7 \mid heart) = \frac{1}{13} = \frac{4}{52} = P(getting\ a\ 7)$

**P35. a.** Let *O* denote the event that the person has type O blood and *not O* denote that the person does not have type O blood.

| | | Second Person | | |
|---|---|---|---|---|
| | | **O** | **Not O** | **Total** |
| **First Person** | **O** | 0.1764 | 0.2436 | **0.4200** |
| | **Not O** | 0.2436 | 0.3364 | **0.5800** |
| | **Total** | **0.4200** | **0.5800** | **1.0000** |

**b.** $P(\text{exactly one of the people has type } O) = 0.2436 + 0.2436 = 0.4872$

**c.**




**P36. a.** P(*at least one has type O blood*) = 1 – P(*none of them have type O blood*)
$= 1 - 0.58^{10} \approx 0.9957$

**b.** P(*at least one doesn't have type O blood*) = 1 – P(*all of them have type O blood*)
$= 1 - 0.42^{10} \approx 0.9998$

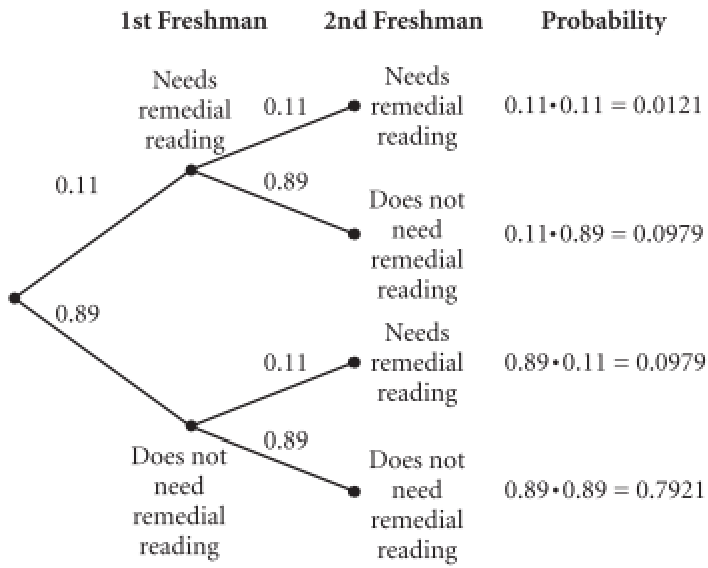
**P37. a.** P(both freshmen must take remedial reading) = P(first freshman must take remedial reading) · P(second freshman must take remedial reading) = 0.11 · 0.11 = 0.0121

**b.**

| | | First Freshman: Needs Remedial Reading | First Freshman: Doesn't Need Remedial Reading | Total |
|---|---|---|---|---|
| Second Freshman | Needs Remedial Reading | 0.0121 | 0.0979 | **0.11** |
| | Doesn't Need Remedial Reading | 0.0979 | 0.7921 | **0.89** |
| | **Total** | **0.11** | **0.89** | **1.00** |

P(*exactly one freshman needs remedial reading*) =
P(*first freshman needs remedial reading* and
*second freshman does not need remedial reading*) +
P(*second freshman needs remedial reading* and
*first freshman does not need remedial reading*) = 0.0979 + 0.0979 = 0.1958.

**c.**




Follow the two branches that represent one needs remedial reading and the other does not, and add the probabilities for those two branches. 0.0979 + 0.0979 = 0.1958.

# Exercise Solutions

**E1. a.** There are 16 possible outcomes:

| 4 Starbucks | 3 Starbucks | 2 Starbucks | 1 Starbucks | No Starbucks |
|---|---|---|---|---|
| SSSS | MSSS<br>SMSS<br>SSMS<br>SSSM | MMSS<br>MSMS<br>MSSM<br>SMMS<br>SMSM<br>SSMM | MMMS<br>MMSM<br>MSMM<br>SMMM | MMMM |

**b.** Since the sample space contains 16 elements, the probability that all 4 people choose Starbucks is $\frac{1}{16}$.

**c.** The probability that all four select Starbucks just by chance is small (0.0625), but not small enough to be convincing. (From Chapter 1, to be convinced that the tasters weren't just guessing, the probability should be less than 0.05.)

**E2. a.** There are 32 possible outcomes. Refer to the table in practice problem **P2(a)** – it is identical with the exception that H and T are now replaced by M and F, respectively.

**b.** Under the assumption that 85% of those injured are male, it is more likely to get a male than a female. So, these outcomes are not equally likely.

**c.** MMMMM is the most likely outcome and FFFFF is the least likely outcome.

**E3. a.** $\frac{30}{36}$ **b.** $\frac{4}{36}$ **c.** $\frac{8}{36}$ **d.** $\frac{6}{36}$

**e.** $\frac{11}{36}$ **f.** $\frac{1}{36}$ **g.** $\frac{9}{36}$ **h.** $\frac{3}{36}$

**i.** $\frac{2}{36}$; these outcomes are (6, 1) and (1, 6).

**E4. a.** $4 \cdot 4$, or 16

**b.**

| | | Second Roll | | | |
|---|---|---|---|---|---|
| | | **1** | **2** | **3** | **4** |
| | **1** | 1, 1 | 1, 2 | 1, 3 | 1, 4 |
| **First Roll** | **2** | 2, 1 | 2, 2 | 2, 3 | 2, 4 |
| | **3** | 3, 1 | 3, 2 | 3, 3 | 3, 4 |
| | **4** | 4, 1 | 4, 2 | 4, 3 | 4, 4 |

**c.** $\frac{4}{16}$ **d.** $\frac{1}{16}$ **e.** $\frac{3}{16}$

**E5. a.** Disjoint and complete; the probabilities for 0, 1, and 2 fours are $\frac{9}{16}$, $\frac{6}{16}$, and $\frac{1}{16}$, respectively.

**b.** Not disjoint but complete; *the first roll is a four* and *the second roll is a four* can occur on the same pair of rolls.
**c.** Disjoint but not complete; only outcomes with fours are included.
**d.** Disjoint but not complete; *the sum equal to 2* is not included.
**e.** Disjoint but not complete; *the second die is a four* is not included.

**E6. a.** Yes, the list is complete.

**b.** No, the outcomes are not disjoint because *heads on second flip* and *heads on first flip* can happen in the same pair of flips.

**c.** No, the event *heads on second flip* has the same probability ($\frac{1}{2}$) as *heads on first flip,* but both are more likely than *heads on neither flip,* which has probability $\frac{1}{4}$.

**d.** No. The probability of getting at least one head is $\frac{3}{4}$, but d'Alembert's sample space gives $\frac{2}{3}$.

**E7. a.** The first two rolls were not doubles, but the third roll was.

**b.** Six rolls were doubles. There were six places where the proportion went up.

**c.** 6/50 or 0.12.

**d.** As the number of rolls increases, one additional roll has a smaller effect on the cumulative proportion. To illustrate this, say you have rolled twice and one of the rolls was a double. The cumulative proportion is 0.5. The next roll will result in a cumulative proportion of 2/3 if the result is another double or 1/3 if the result is not a double. Either way it is a change of about 0.167. Now imagine there have been 100 rolls and 50 of them were doubles for a cumulative proportion of 0.5. The next roll would result in a cumulative proportion of 50/101 or 51/101. Either way this is a much smaller change in the cumulative proportion, which would result in a shorter segment.

**E8. a.** 227/500 = 0.454

**b.** According to the Law of Large Numbers, since Jack has conducted more trials, his estimate would be expected to be closer to the true probability.

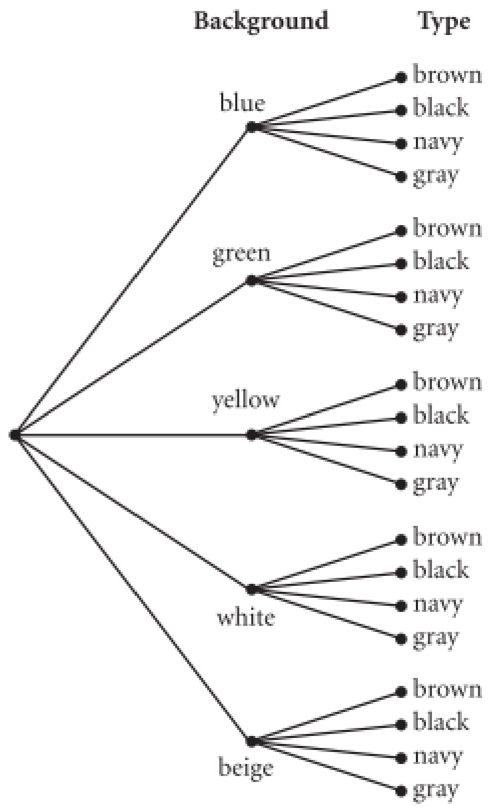
**c.** Yes. As the number of flips increases, the percentage of heads gets closer to 50%.

**d.** The number of heads gets further from half the number of flips. In the case of 100 flips, the result is only six away from the 50 heads that would be expected. After 10,000 flips, there were 140 more heads than expected. This is evidence against a common misperception about the Law of Large Numbers. Many people think that if, for example, heads turns up five times in a row, then tails is more likely to turn up next. As if the coin is somehow self-correcting. If that were the case then we would expect the number of heads to get closer to half the number of flips, not further.

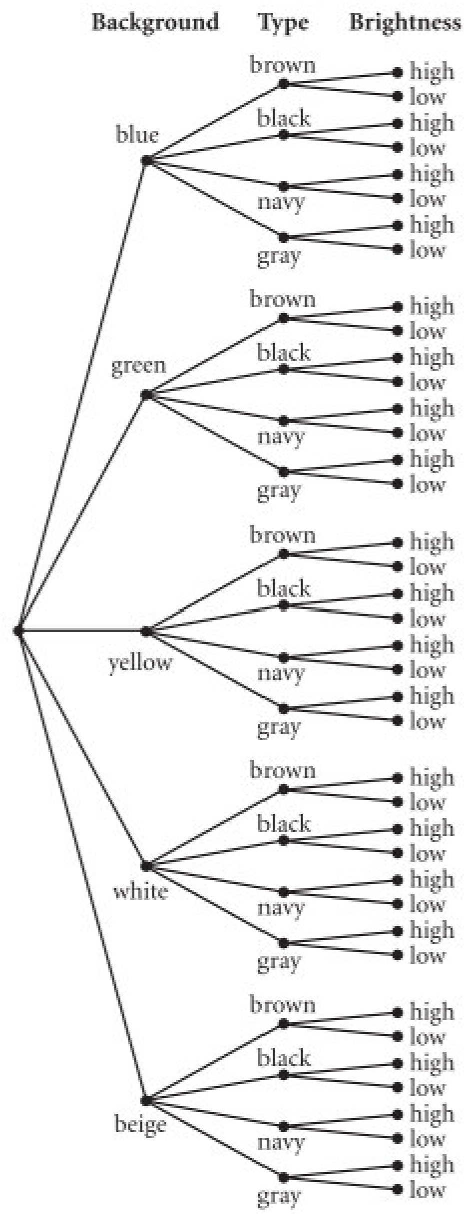
**E9. a.** Five different backgrounds and four different types give $5 \cdot 4$, or 20, possible treatments.

**c.** With two levels of brightness added, the number of possible treatments is now $5 \cdot 4 \cdot 2$, or 40.

**b.**




**d.**




**E10.** **a.** 6 · 4, or 24

**b.**

**Shirts**

| | | Blue | Green | Red | Yellow | White (Long) | White (Short) |
|---|---|---|---|---|---|---|---|
| **Pants** | **Brown** | Brown, Blue | Brown, Green | Brown, Red | Brown, Yellow | Brown, White (Long) | Brown, White (Short) |
| | **Black** | Black, Blue | Black, Green | Black, Red | Black, Yellow | Black, White (Long) | Black, White (Short) |
| | **Blue** | Blue, Blue | Blue, Green | Blue, Red | Blue, Yellow | Blue, White (Long) | Blue, White (Short) |
| | **Gray** | Gray, Blue | Gray, Green | Gray, Red | Gray, Yellow | Gray, White (Long) | Gray, White (Short) |

**c.** $P(\textit{a white shirt}) = \frac{8}{24}, \frac{2}{6}$ or $\frac{1}{3}$ $P(\textit{the gray pants}) = \frac{6}{24}$, or $\frac{1}{4}$

*P*(*a white shirt* and *the gray pants*) = $\frac{2}{24}$, or $\frac{1}{12}$ *P*(*a white shirt* or *the gray pants*) = $\frac{12}{24}$, or $\frac{1}{2}$

**E11. a.** Yes. Both candidates have an equal chance of selecting the white marble.

**b.** If they both draw a white marble on the same round they would tie again. The probability of this happening on any given round would be 1/9 · 1/9 = 1/81.

**c.** Adding more non-white marbles to the bags would reduce the chance of a tie, but would also increase the expected number of draws before a winner is declared.

**E12. a.** There is a 1/27 chance of your favorite song being selected, so the probability of it happening twice in a row is 1/27 · 1/27 = 1/729.

**b.** Here any song can be selected first, then there is a 1/27 chance that the same song will be selected a second time.

**E13. a.** There are $2^6$, or 64, equally likely outcomes; the probability of getting heads all six times is $\frac{1}{64}$.

**b.** There are 6 · 6 · 6 · 6 · 6 · 6, or 46,656, equally likely outcomes; the probability of getting a 3 all six times is $\frac{1}{46,656}$.

**c.** No. There are two choices for each person (school, no school) and so $2^{1200}$ possible outcomes. They are not equally likely (far fewer people are in school than are out of school), so you cannot find the probability without further information.

**E14.** The number of three-digit numbers is 3 · 3 · 3, or 27. The number of outcomes less than 250 include all those that start with 1 and all those that start with 21 and 22. If the number starts with 1, there are 3 numbers to pick from for the second digit and 3 for the third digit, so there are 3 · 3, or 9, three-digit numbers that start with 1. If the number starts with 21 it can end with 1, 2, or 7. The same is true if the number starts with 22. So there are 6 three-digit numbers that start with 21 or 22. Therefore, the probability is $\frac{15}{27}$.

**E15.** a and c

**E16.** b and d. Students could argue that the events in d are not mutually exclusive, saying it is possible that a 95 year old has a living mother.

**E17. a.** The categories are not disjoint because the numbers in the three categories add up to more than the total number of people who fish in the US. Some people fish in more than one category of water. The categories are complete because the two kinds of bodies of water are fresh and salt water. The Great Lakes are separated out as a separate category of fresh water.

**b.** Yes. $\frac{7.7}{30} \approx 0.257$

**c.** No because you don't know how many people who fish in the Great Lakes also fish in other freshwater.

**d.** P(fresh or salt) = P(fresh) + P(salt) – P(fresh and salt)

30,000,000 / 30,000,000 = 26,400,000 / 30,000,000 + 7,700,000 / 30,000,000 – P(fresh and salt)

P(fresh and salt) = (34,100,000 – 30,000,000)/30,000,000 = 0.1367

**E18. a.** The complete disjoint categories are the last four: received full amount, received partial amount, did not receive payments, and not supposed to receive payments in 2005. A revised table with complete and disjoint categories is as follows:

| Child support received by custodial parents | Numbers (in thousands) |
|---|---|
| With child support agreement or award, supposed to receive payments in 2005 and actually received full amount | 3,192 |
| With child support agreement or award, supposed to receive payments in 2005 and actually received partial payment | 2,067 |
| With child support agreement or award, supposed to receive payments in 2005 and didn't receive any payments | 1,550 |
| With child support agreement or award, but not supposed to receive payments in 2005 | 993 |
| Child support not awarded | 5,803 |
| **Total custodial parents with children under age 21** | **13,605** |

**b.** Because the categories *received full amount* and *received partial amount* are disjoint, we can add cases to obtain

$$P(\text{full amount or partial amount}) = \frac{3192 + 2067}{13{,}605} \approx 0.387 .$$

**E19. a.**

| Receives financial aid | Owns a Laptop? Yes | No | Total |
|---|---|---|---|
| Yes | 26 | 46 | 72 |
| No | 112 | 32 | 144 |
| Total | 138 | 78 | 216 |

**b.** The probability that a randomly selected student doesn't receive financial aid or doesn't own a laptop is (144 + 78 – 32) / 216 ≈ 88.0%. Alternately, the complement of "doesn't receive financial aid or doesn't own a laptop" is "receives financial aid and has a laptop. 1 – 26 / 216 ≈ 88.0%.

**E20. a.**

| | Driver Violated a Traffic Law | Driver Didn't Violate a Traffic Law | Total |
|---|---|---|---|
| Fatality | 521 | 316 | 837 |
| No Fatality | 144,767 | 8,303 | 153,070 |
| Total | 145,288 | 8,619 | 153,907 |

**b.** 521 / 153,907 $\approx$ 0.0034 of the crashes involved a fatality and a driver who violated a traffic law.

**c.** 1 – 8,303 / 153,907 $\approx$ 0.946 of the crashes involved a fatality or a driver who violated a traffic law.

**d.** The proportion of fatal crashes that involved a driver who violated traffic law is $\frac{521}{837} \approx 0.622$. This is much lower than the proportion of nonfatal accidents involving a driver who violated traffic law, namely $\frac{144,767}{153,070} \approx 0.946$.

**E21. a.**

Genetic Mutation A? / Genetic Mutation B?

| | Yes | No | Total |
|---|---|---|---|
| Yes | 55 | 25 | 80 |
| No | 15 | 5 | 20 |
| Total | 70 | 30 | 100 |

The proportion who have A or have B is 0.55 + 0.25 + 0.15 = 0.95, or 95%

**b.** $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 0.80 + 0.70 - 0.55 = 0.95$.

**E22.** Let $x =$ P(*abandoned* and *woman*). We are given the non-boldfaced information in the following table:

| | Abandoned (A) | Stayed Together | |
|---|---|---|---|
| Woman (W) | **0.11** | 0.42 | 0.53 |
| Man | **0.01** | **0.46** | **0.47** |
| | 0.12 | **0.88** | **1.00** |

In order to fill in the rest of the table (the boldfaced quantities), we need to compute P(A and W). Using the addition law yields:

$$P(A \text{ or } W) = P(A) + P(W) - P(A \text{ and } W)$$
$$0.54 = 0.12 + 0.53 - P(A \text{ and } W)$$
$$0.11 = P(A \text{ and } W)$$

**E23.**

| | Employed | Unemployed | Not in Labor Force | Total |
|---|---|---|---|---|
| High School Graduates | 5,142,530 | 796,730 | 1,303,740 | **7,243,000** |
| High School Dropouts | 1,995,980 | 527,240 | 1,242,780 | **3,766,000** |
| Total | **7,138,510** | **1,323,970** | **2,546,520** | **11,009,000** |

The probability that a randomly selected person aged 16 to 24 who isn't enrolled in college or high school is employed is 7,138,510 / 11,009,000 ≈ 0.648.

**E24.** Letting *D* denote Democrat, *R* Republican, and *A* approve, we have:

$$P(R \text{ or } A) = P(R) + P(A) - P(R \text{ and } A) = \frac{640}{1500} + \frac{937}{1500} - \frac{449}{1500} = \frac{1128}{1500} = 0.752.$$

Alternatively, make a table:

| | Democrat | Republican | Total |
|---|---|---|---|
| Approve | 488 | 449 | **937** |
| Not Approve | 372 | 191 | **563** |
| Total | **860** | **640** | **1500** |

Observe that

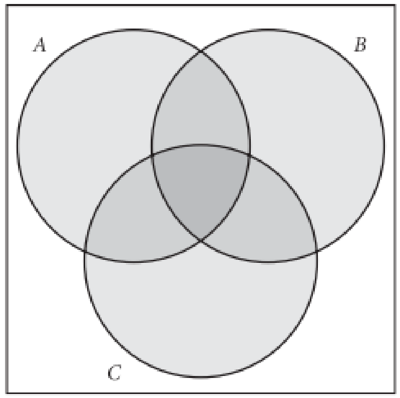
$$P(R \text{ or } A) = \frac{(488 + 449 + 191)}{1500} = \frac{1128}{1500} = 0.752.$$

**E25.** Jill has not computed correctly because it is not true that these two events are mutually exclusive. It is common to use the language that "two events are mutually exclusive if they can not happen at the same time." As you can see from this example, language can be misleading. These two events are not mutually exclusive even though they can not happen at the same instance in time. However, they can happen in the same situation. The first flip can be a head and the second flip can be a head, or *HH,* so the two events are not mutually exclusive.

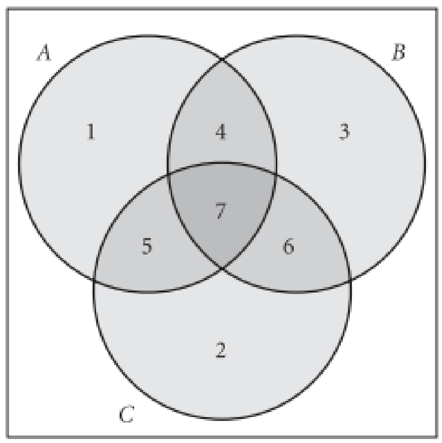
**E26.** Jill has computed correctly using the right definition of mutually exclusive. *HT* and *TH* are mutually exclusive because one event involves getting a head on the first flip of the coin and the other event involves getting a tail. These cannot both happen in the same situation.

**E27.** Each of the statements *P*(*A* and *B*) = 0, *P*(*B* and *C*) = 0, and *P*(*A* and *C*) = 0 must be true in order for *P*(*A* or *B* or *C*) = *P*(*A*) + *P*(*B*) + *P*(*C*).

(Note that technically we also need *P*(*A* and *B* and *C*) = 0, but this is implied by each of these three statements.) To understand why this is true, examine this Venn diagram.




**E28. a.**




**b.** Observe that
P(A) = 1+4+5+7, P(B) = 3+4+6+7, P(C) = 2+5+6+7, $P(A \cup B)$ = 4+7,
$P(B \cup C)$ = 6+7, $P(A \cup C)$ = 5+7, $P(A \cup B \cup C)$ = 7.

Now, using this information yields

$$P(A) + P(B) + P(C) - P(A \cup B) - P(B \cup C) - P(A \cup C) + P(A \cup B \cup C) =$$
$$1+4+5+7 + 3+4+6+7 + 2+5+6+7 - (4+7) - (6+7) - (5+7) + 7 =$$
$$1+2+3+4+5+6+7.$$

Hence, the rule is:

P( A or B or C) = P(A) + P(B) + P(C) – P(A and B) – P(B and C) – P(A and C) + P(A and B and C)

**E29. a.** $\frac{60,838}{232,556} \approx 0.262$

**b.** $\frac{7,798}{37,490} \approx 0.208$

**c.** $\frac{7,798}{60,838} \approx 0.128$

**d.** $\frac{29,692 + 30,888}{171,718} \approx 0.353$

**e.** Age 35 to 44 has $\frac{12,902}{42,301} \approx 0.305$, or 30.5% volunteers. Age 16 to 24 has 20.8% volunteers.

**E30.**

| | | Carbolic Acid Used? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Survived? | Yes | 34 | 19 | 53 |
| | No | 6 | 16 | 22 |
| | Total | 40 | 35 | 75 |

**a.** $\frac{6}{40}$ **b.** $\frac{6}{22}$ **c.** $\frac{6}{75}$ **d.** $\frac{56}{75}$

**E31. a.**

| | | Receives additional medical attention? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Age | Under 90 | 0.44 | 0.48 | 0.92 |
| | 90 and older | 0.08 | 0 | 0.08 |
| | Total | 0.52 | 0.48 | 1.00 |

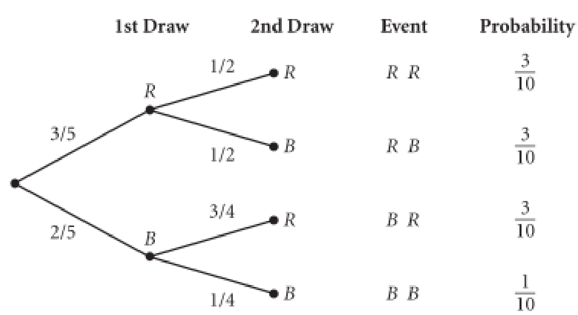
**b.** P(90 or older | receives additional medical attention) = $\frac{0.08}{0.52} \approx 0.15$

**c.** P(90 or older and does not receive additional medical attention) = 0

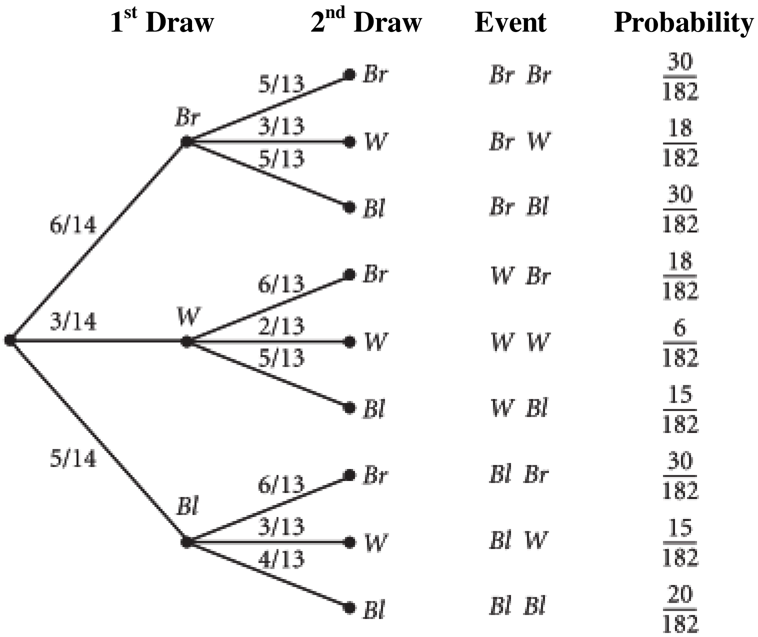
**E32.** No, you need one additional piece of (joint) information in order to complete the table. Information on any one cell will do the job.

**E33. a.** $\frac{3}{5}$ **b.** $\frac{2}{5}$ **c.** $\frac{1}{2}$ **d.** $\frac{1}{2}$ **e.** $\frac{3}{4}$ **f.** $\frac{1}{4}$

Here is the tree diagram for drawing marbles:



**E34. a.**




**b.** 3/14
**c.** 2/13
**d.** 3/13
**e.** 6/14 · 5/13 + 3/14 · 2/13 + 5/14 · 4/13 ≈ 0.308

**E35.** **a.** 4/52 · 4/52 ≈ 0.0059
**b.** 4/52 · 3/51 ≈ 0.0045
**c.** 4/52 · 4/52 ≈ 0.0059
**d.** 4/52 · 4/51 ≈ 0.0060
**e.** 1 · 13/52 (The first card can be anything, and the second card can be any of the 13 cards of the same suit.)
**f.** 1 · 12/51 ≈ 0.2353

**E36.** Because 0.36 · 0.62 = P(*interested*) · P(< *age 25* | *interested*) = P(*interested* and < *age 25*), the question is, if you select an adult at random from the United States, what is the probability that the person wants to go to Mars and is under 25 years old?

**E37.** This problem is easiest to understand if you first make a table.

| | | Rain Predicted? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Rain? | Yes | 6 | 18 | 24 |
| | No | 4 | 72 | 76 |
| | Total | 10 | 90 | 100% |

The desired probability is 0.24, or 24%.

**E38.** For a typical set of 100 games, the won-lost record would look approximately like the following table. Note that the 36% is the conditional percentage of wins given that the team is facing a left-handed starting pitcher or $P(\textit{win} \mid \textit{left-handed pitcher})$. This means that, in order to calculate the cell for *win* and *left-handed pitcher* you must calculate 36% of the 18 games against left-handed pitchers = 6.48.

| | Right-Handed Pitcher | Left-Handed Pitcher | Total |
|---|---|---|---|
| **Won** | 46.52 | 6.48 | **53** |
| **Lost** | 35.48 | 11.52 | **47** |
| **Total** | **82** | **18** | **100** |

The percentage of wins given a right-handed starting pitcher is $\frac{46.52}{82} \approx 0.57$, or 57%.

**E39. a.** Here are the expected results for 20 contaminated samples:

| | | Technician's Decision | | |
|---|---|---|---|---|
| | | Positive | Negative | Total |
| Contamination | Present | 18 | 2 | **20** |
| | Absent | 8 | 72 | **80** |
| | Total | **26** | **74** | **100** |

**b.** The false positive rate is $\frac{8}{26} \approx 0.31$. Equivalently, we say PPV = $\frac{18}{26} \approx 0.69$.
**c.** The false negative rate is $\frac{2}{74} \approx 0.03$. Equivalently, we say NPV = $\frac{72}{74} \approx 0.97$.
**d.** With 50 contaminated samples out of 100, the expected results change as shown here.

| | | Test Result | | |
|---|---|---|---|---|
| | | Positive | Negative | Total |
| Contamination | Present | 45 | 5 | **50** |
| | Absent | 5 | 45 | **50** |
| | Total | **50** | **50** | **100** |

The false positive rate is $\frac{5}{50} = 0.10$, or equivalently PPV = 0.90. The false negative rate is $\frac{5}{50} = 0.10$, or equivalently NPV = 0.90. It is important to note that the false positive rate goes down when the population being tested has a higher proportion of contaminated cases.

**E40. a.** For Officiousville, the table is

| | | Test Result | | |
|---|---|---|---|---|
| | | Positive | Negative | Total |
| Disease | Present | 450 | 50 | **500** |
| | Absent | 50 | 450 | **500** |
| | Total | **500** | **500** | **1000** |

The false positive and false negative rates are both $\frac{50}{500} = \frac{1}{10} = 0.10$, or equivalently both PPV and NPV = 0.90.

**b.** For Mellowville, the table is

| | | Test Result | | |
|---|---|---|---|---|
| | | **Positive** | **Negative** | **Total** |
| **Disease** | **Present** | 9 | 1 | **10** |
| | **Absent** | 99 | 891 | **990** |
| | **Total** | **108** | **892** | **1000** |

The false positive rate is $\frac{99}{108} \approx 0.92$, or equivalently PPV = 0.08; the false negative rate is $\frac{1}{892} \approx 0.001$, or equivalently NPV = 0.999.

**E41.** **a.** No. The false positive rate is computed using

$$\frac{\text{number of false positive tests}}{\text{number of positive tests}}$$

You are given only the total number of tests, not the total number of positive tests. If all of the 28,436 tests were positive, the false positive rate would be $\frac{107}{28,436} \approx 0.0038$ or less than half of one percent. If only 107 tests were positive, the false positive rate would be 107/107 = 100%.

**b.** This is a new test so it is unlikely that false negatives have been reported yet. People who had negative tests were told that they tested negative for H.I.V. and so may not know that the test gave a false result until symptoms of AIDS appear. On the other hand, false positives will be reported right away because anyone who tests positive on this test will be given further studies to determine the extent of the H.I.V. infection.

**c.** Let $x$ be the number of false negatives. Then

$$\frac{x+107}{28,436} = 0.01$$

$$x = 177.36 \approx 177$$

**E42.** Answers will vary. For example, consider the following table for a screening test that is 90% accurate (both sensitivity and specificity are 0.9).

| | | Test Result | | |
|---|---|---|---|---|
| | | **Positive** | **Negative** | **Total** |
| **Disease** | **Present** | 720 | 80 | **800** |
| | **Absent** | 2 | 18 | **20** |
| | **Total** | **722** | **98** | **820** |

The false positive rate is $\frac{2}{722} \approx 0.003$. The false negative rate is $\frac{80}{98} \approx 0.816$.

**E43.** A table (or a tree diagram) might help here.

| | | Type of Coin | | |
|---|---|---|---|---|
| | | Fair Coin | Two-Headed | Total |
| Flip Result | Heads | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{3}{4}$ |
| | Tails | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ |
| | Total | $\frac{1}{2}$ | $\frac{1}{2}$ | 1 |

**a.** $\frac{1}{2}$ **b.** 1 **c.** $\frac{1}{3}$ **d.** $\frac{2}{3}$ **e.** 1 **f.** 0

**E44.** The first statement is false, and the second is true. To see this, consider a two-way table such as this one.

| | *B* | *Not B* | Total |
|---|---|---|---|
| *A* | 72 | 8 | 80 |
| *Not A* | 2 | 18 | 20 |
| Total | 74 | 26 | 100 |

$P(A \mid B) + P(A \mid not\ B) = \frac{72}{74} + \frac{8}{26}$, which exceeds 1, so this can not be the probability of *A*.
It is true that $P(A$ and $B) + P(A$ and $not\ B) = \frac{72}{100} + \frac{8}{100} = \frac{80}{100} = P(A)$.

**E45. a.** $P(doubles) = \frac{1}{6}$; $P(doubles \mid sum\ of\ 8) = \frac{1}{5}$. Not independent.

**b.** $P(sum\ of\ 8) = \frac{5}{36}$; $P(sum\ of\ 8 \mid 2\ on\ first) = \frac{1}{6}$. Not independent.

**c.** $P(sum\ of\ 7) = \frac{1}{6}$; $P(sum\ of\ 7 \mid 1\ on\ first) = \frac{1}{6}$. Independent.

**d.** $P(doubles) = \frac{1}{6}$; $P(doubles \mid sum\ of\ 7) = 0$. Not independent.

**e.** $P(1\ on\ first) = \frac{1}{6}$; $P(\text{1 on first} \mid \text{1 on second}) = \frac{1}{6}$. Independent.

**E46. a.** These are not independent because $P(test\ positive \mid relative) > P(test\ positive)$.

**b.** These are independent because we expect that

$$P(test\ is\ positive \mid last\ digit\ is\ a\ 3) = P(test\ is\ positive).$$

**c.** These are not independent because the states where the highest elevation isn't very large tend to lie east of the Mississippi and so,

$$P(east\ of\ Miss. \mid over\ 8000\ feet) < P(east\ of\ Miss.)$$

and

$$P(over\ 8000\ feet \mid east\ of\ Miss.) < P(over\ 8000\ feet).$$

In fact, both conditional probabilities are 0.

**E47. a.** *P*(*left-handed*) = 12 / 100 *P*(*left eye dominant*) = 37 / 100
*P*(*left eye dominant* | *left-handed*) = 6/12 *P*(*left-handed* | *left eye dominant*) = 6 / 37

**b.** No. *P*(*left eye dominant*) $\neq$ P(*left eye dominant* | *left-handed*).
**c.** No. P(*left eye dominant* | *left-handed*) $\neq$ 0.

**E48. a.** *P*(*admitted*) = 870 / 1725 ≈ 0.504 *P*(*admitted* | *woman*) = 220 / 483 ≈ 0.455
*P*(*admitted* | *man*) = 650 / 1242 ≈ 0.52

**b.** No, a randomly selected woman is less likely to be admitted than a person randomly selected from all applicants.

**c.** No, P(*admitted* | *woman*) $\neq$ 0

**d.** According to the data, women are less likely to be accepted than men. There may be enough evidence to ask the university for an explanation.

**e.** No. In Program A, P(*admitted*) = 601 / 933 ≈ 0.644. P(*admitted* | *woman*) = 89 / 108 ≈ 0.82. P(*admitted* | *man*) = 512 / 825 ≈ 0.62. Women are more likely than a person randomly selected from all applicants to be accepted to Program A.

**f.** No. In Program B, P(*admitted*) = 269 / 792 ≈ 0.340. P(*admitted* | *woman*) = 131 / 375 ≈ 0.349. P(*admitted* | *man*) = 138 / 417 ≈ 0.331. Again, women are more likely than a person randomly selected from all applicants to be accepted to this program.

**g.** The apparent paradox is that, while women are more likely to be accepted into either program, when the data are combined the opposite is true. Sometimes the direction of comparison in subgroups is reversed when the subgroups are combined. This phenomenon is known as Simpson's paradox.

**E49. a.** *BB, BG, GB,* and *GG*
**b.** two boys; two girls
**c.** (0.51)(0.51), or 0.2601, assuming births are independent
**d.** They may mean that their probability of getting a girl is higher than the population percentage for girls, and higher than the percentage for boys, if girls "run in the family." Under these conditions, *GG* would be the event with the highest probability.

**E50. a.** Let Y = Believes in astrology, N = Does not believe in astrology.
There are 16 possible outcomes. The listing is exactly as in E1(a), with the exception that S and M are replaced by Y and N, respectively.

**b.** NNNN is the most likely outcome and YYYY is the least likely outcome.

**c.** P(NNNN) = (0.75)(0.75)(0.75)(0.75) = 0.316. People must be chosen in such a way that their opinions are not influenced by any others' in the sample. That is, the trials must be independent.

**E51. a.** Because these two events are independent, we can use the Multiplication Rule for Independent Events:

$$P(type\ O \text{ and } Rh+) = (0.42)(0.05) = 0.021$$

**b.** $P(type\ O \text{ or } Rh+) = P(type\ O) + P(Rh+) - P(type\ O \text{ and } Rh+) = 0.42 + 0.05 - 0.021 = 0.449$

**c.**

| | | Rh-Positive? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Type O? | Yes | 2.1 | 39.9 | 42 |
| | No | 2.9 | 55.1 | 58 |
| | Total | 5 | 95 | 100% |

**E52.** A two-way table will help determine which questions can be answered.

| | | Schizophrenic? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Employed? | Yes | 0.003 | | |
| | No | 0.007 | | |
| | Total | **0.01** | **0.99** | **1.00** |

**a.** 0.01
**b.** This cannot be answered based on the given information.
**c.** 0.003
**d.** This cannot be answered based on the given information.
**e.** More than 70%. Probably only slightly more than 70% or the rate would have been reported differently.
**f.** This cannot be answered based on the given information.

**E53. a.** Assuming that all batches are equally likely to be chosen, we seek P (J and J) = 0.25.
**b.** Suggest that each highway get 1 batch of cement H and 1 batch of cement J. Since the two sections comprising the highway are identical, using one type of cement on the two sections will yield results as to how each type of cement performs on two different highways.

**E54.** P(*all juniors*) = $6/11 \cdot 5/10 \cdot 4/9 \approx 0.12$.
P(*all seniors*) = $5/11 \cdot 4/10 \cdot 3/9 \approx 0.06$.
P(*all juniors* or *all seniors*) $\approx 0.12 + 0.06 = 0.18$.

**E55. a.** P(*all 12 require remedial reading*) = $0.11^{12} = 3 \cdot 10^{-12}$.
**b.** P(*at least one requires remedial reading*) = 1 – P(*none require remedial reading*)
$= 1 - 0.89^{12} = 0.753$.

**E56. a.** P(*none are schizophrenic*) = $0.99^{50} \approx 0.605$
**b.** P(*at least one is schizophrenic*) = 1 – P(*none are schizophrenic*) $\approx 1 - 0.605 = 0.395$

**E57.** The results of the surgeries must be assumed to be independent in order to find the probability both surgeries will fail, with the given information. This typically is not a reasonable assumption, because the two surgeries are to be performed on the same person at the same time. If the surgeries can be assumed to be independent, as perhaps with minor surgery performed on different parts of the body, $P(A\ fails) = 0.15$ and $P(B\ fails) = 0.10$ and the probability that they both fail is the product 0.015.

**E58.** Jill's probability is for the event *4 on first* and *not 4 on the second.* She forgot to account for the event *not 4 on first* and *4 on second.* The correct probability is $\frac{10}{36}$.

**E59.** A two-way table will help. Begin by filling in the marginal totals for left and right-handed pitchers. Then calculate the probabilities in the cells under "Yes" as shown below. After that, adding and subtracting will allow you to fill in the rest of the table.

| Pitcher? | Successful Hit? Yes | No | Total |
|---|---|---|---|
| Left-handed | $0.15 \cdot 0.20 = 0.03$ | 0.17 | **0.20** |
| Right-handed | $0.23 \cdot 0.80 = 0.184$ | 0.616 | **0.80** |
| Total | **0.214** | **0.786** | **1.00** |

He hits successfully 21.4% of the time.

**E60.** Dave's arithmetic is not too bad. He is right that 2 percent means odds of 98 to 2, or 49 to 1, against an attack occurring on his rug if the dog selected a spot at random. However, if four attacks were to occur, each at a randomly selected spot in the house and each independent of the other, the probability all four would be on the rug is $(.02)^4 = (\frac{1}{50})^4 = \frac{1}{6,250,000}$. This is one chance in about six million (or odds against of about 6,249,999 to 1), not one chance in five million. (Dave probably got one chance in five million by computing $49^4$, or 5,764,801.)

However, Dave's logic is not too good because his computation assumed that the attacks were independent. Surely a sick dog would tend to be sick repeatedly in the same location rather than randomly wandering around the house.

**E61.** Let's call these three conditions I, II, and III.

I: $P(A) = P(A \mid B)$ II: $P(B) = P(B \mid A)$ III: $P(A \text{ and } B) = P(A) \cdot P(B)$

<u>The proof that I ⇒ II is as follows</u>:

Suppose $P(A \mid B) = P(A)$. Then by the definition of conditional probability,

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{P(B) \cdot P(A \mid B)}{P(A)} = \frac{P(B) \cdot P(A)}{P(A)} = P(B)$$

By switching $A$ and $B$, this proof also shows that II ⇒ I.

<u>The proof that I ⇒ III is as follows</u>:

First write the definition of conditional probability:

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

Because I is assumed to be true, substitute $P(A)$ for $P(A \mid B)$ to get

$$P(A) = \frac{P(A \text{ and } B)}{P(B)}$$

Multiplying both sides by $P(B)$ gives

$$P(A) \cdot P(B) = P(A \text{ and } B)$$

which is III.

<u>The proof that III ⇒ I is as follows</u>:
Because III is true, we have

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

Dividing both sides by $P(B)$ gives

$$P(A) = \frac{P(A \text{ and } B)}{P(B)}$$

But from the definition of conditional probability, we know that

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

Comparing the last two equations, we see that $P(A) = P(A \mid B)$. So I is true.

Because we have shown that II ⇔ I and I ⇔ III, we get by transitivity that II ⇔III, and the three conditions are equivalent.

**E62.** We are given that A and B are independent. The multiplication rule for independent events tells us $P(A \text{ and } B) = P(A) \cdot P(B)$.

As in E44, we see from any two-way table that

$$P(A \text{ and } \textit{not } B) + P(A \text{ and } B) = P(A).$$

So,

$$\begin{aligned} P(A \text{ and } \textit{not } B) &= P(A) - P(A \text{ and } B) \\ &= P(A) - P(A) \cdot P(B) \\ &= P(A) \cdot [1 - P(B)] \\ &= P(A) \cdot P(\textit{not } B) \end{aligned}$$

Therefore, from the multiplication rule, *A* and *not B* are independent.

**E63. a.** Think of a table with *A* and *B* intersecting. The symbol $\overline{A}$ denotes *not A*.

| | $B$ | $\overline{B}$ |
|---|---|---|
| $A$ | | |
| $\overline{A}$ | | |

From the table, $P(A) = P(B \text{ and } A) + P(\overline{B} \text{ and } A)$, which, using the Multiplication Rule, is equal to $P(B) \cdot P(A \mid B) + P(\overline{B}) \cdot P(A \mid \overline{B})$.

**b.** Use this to find *P*(*Dodgers win*):

$$P(win) = P(win \mid day) \cdot P(day) + P(win \mid not\ day) \cdot P(not\ day) = \frac{11}{21} \cdot \frac{21}{78} + \frac{30}{57} \cdot \frac{57}{78} = \frac{41}{78}$$

**c.** $P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)} = \frac{P(B) \cdot P(A|B)}{P(A)}$

But from part a, $P(A) = P(A \mid B) \cdot P(B) + P(A \mid \overline{B}) \cdot P(\overline{B})$.

**d.**

$$P(win \mid day) = \frac{P(day \mid win) \cdot P(win)}{P(day \mid win) \cdot P(win) + P(day \mid not\ win) \cdot P(not\ win)} = \frac{\frac{11}{41} \cdot \frac{41}{78}}{\frac{11}{41} \cdot \frac{41}{78} + \frac{10}{37} \cdot \frac{37}{78}} = \frac{11}{21}$$

**E64. a.** Using *I* for *Test I is used, II* for *Test II is used,* and *F* for *false positive indicated,* the problem gives $P(F \mid I) = .10$, $P(F \mid II) = .05$, and $P(I)=0.6$. These values can be used to complete the table.

| | Test I | Test II | Total % |
|---|---|---|---|
| False Positives | 6 | 2 | 8 |
| Other Results | 54 | 38 | 92 |
| Total % | 60 | 40 | 100% |

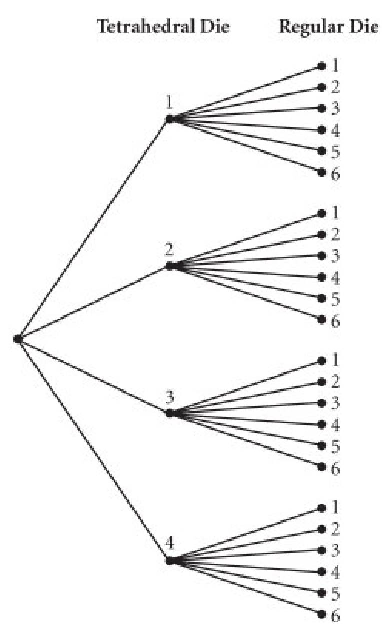
**b.** From the table, $P(I \mid F) = \frac{6}{8} = 0.75$. Also, from the formula, we obtain

$$P(I \mid F) = \frac{P(F \mid I) \cdot P(I)}{P(F \mid I) \cdot P(I) + P(F \mid II) \cdot P(II)} = \frac{(0.1)(0.6)}{(0.1)(0.6) + (0.05)(0.4)} = 0.75.$$

**E65. a.** $4 \cdot 6$, or 24

**b.**

| | | Regular Die | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| Tetrahedral Die | 1 | 1, 1 | 1, 2 | 1, 3 | 1, 4 | 1, 5 | 1, 6 |
| | 2 | 2, 1 | 2, 2 | 2, 3 | 2, 4 | 2, 5 | 2, 6 |
| | 3 | 3, 1 | 3, 2 | 3, 3 | 3, 4 | 3, 5 | 3, 6 |
| | 4 | 4, 1 | 4, 2 | 4, 3 | 4, 4 | 4, 5 | 4, 6 |



**c.** $\frac{4}{24}$

**d.** $\frac{2}{24}$

**e.** These events are not disjoint because you can roll (2, 2). These events are not independent events because *P*(*sum of 4*) = $\frac{3}{24}$, but *P*(*sum of 4* | *doubles*) = $\frac{1}{4}$.

**f.** These events are not disjoint because you can roll (2, 5). These events are independent because *P*(*2 on tetrahedral*) = $\frac{1}{4}$ = *P*(*2 on tetrahedral* | *5 on the regular die*).

**E66.** **a.** The 15 outcomes in bold represent the rolls that enable you to hit your opponent. The probability of such a roll is 15/36.

(1, 1) (1, 2) (1, 3) **(1, 4)** **(1, 5)** (1, 6)
(2, 1) (2, 2) **(2, 3)** (2, 4) **(2, 5)** (2, 6)
(3, 1) **(3, 2)** (3, 3) (3, 4) **(3, 5)** (3, 6)
**(4, 1)** (4, 2) (4, 3) (4, 4) **(4, 5)** (4, 6)
**(5, 1)** **(5, 2)** **(5, 3)** **(5, 4)** **(5, 5)** **(5, 6)**
(6, 1) (6, 2) (6, 3) (6, 4) **(6, 5)** (6, 6)

**b.** Yes, the events *sum of 5* and *5 on either die* are mutually exclusive:

$$P(\textit{sum of } 5 \text{ or } 5 \textit{ on either die}) = P(\textit{sum of } 5) + P(5 \textit{ on first die} \text{ or } 5 \textit{ on second die})$$

$$= \frac{4}{36} + \frac{11}{36} = \frac{15}{36}$$

**E67.** P(*favorite song plays at least once*) = 1 – P(*favorite song does not play)* = $1-(\frac{8}{9})^7 \approx .56$.

**E68.** Marilyn correctly gave the answer that the man has a larger chance of having two boys. All her readers agreed that the probability that the man has two boys is the same as the probability that the younger child is a boy: $\frac{1}{2}$. There are only two possible families for him, and each is equally likely: *BG* and *BB*. But Marilyn's readers disagreed on the probability that the woman has two boys. Many thought it also was $\frac{1}{2}$. However, the possible pairs for the woman are *GB, BG,* and *BB*. We are given no reason to doubt that these are equally likely. So the probability she has two boys is only $\frac{1}{3}$. Some of Marilyn's readers got quite angry with her over this.

**E69.** Since both were observed at random, the events

*man washes his hands*

and

*woman washes her hands*

are independent.

**a.** *P*(*both the man and the woman wash their hands*) = 0.75 · 0.90 = 0.675
**b.** *P*(*neither wash their hands*) = 0.25 · 0.10 = 0.025
**c.** *P(at least one washes their hands*) = 1 – P(*neither wash their hands*) = 0.975.
**d.** *P*(*man washes his hands* | *woman washes her hands*) = P(*man washes his hands*) = 0.75

**E70. a.** P(both believe in astrology) = 0.25 · 0.25 = 0.0625
**b.** P(neither believes in astrology) = 0.75 · 0.75 = 0.5625
**c.** P(at least one believes in astrology) = 1 – P(neither believes in astrology) = 0.4375
**d.** P(the second person believes in astrology | the first person believes in astrology) =
P(the second person believes in astrology) = 0.25

**E71. a.** $P(\textit{two teachers}) = \frac{2}{4} \cdot \frac{1}{3} = \frac{1}{6} \approx .167$

**b.** $P(\textit{two teachers}) = \frac{5{,}000{,}000}{10{,}000{,}000} \cdot \frac{4{,}999{,}999}{9{,}999{,}999} = .249999975$

**c.** The draws in both (a) and (b) are technically dependent, but the large population size in (b) makes independence a reasonable assumption. The probability is essentially (0.5)(0.5) = 0.25 for the large population.

**E72. a.** For jumbo mortgages, we have the following information:

| | Late payment | No late payment | Total |
|---|---|---|---|
| **Excellent credit rating** | 0.05 | 0.40 | 0.45 |
| **Not excellent credit rating** | 0.15 | 0.40 | 0.55 |
| **Total** | 0.20 | 0.80 | 1.00 |

Using the addition law, we have:

P(*no late payment* or *excellent credit rating*) =
P(*no late payment*) + P(*excellent credit rating*) - P(*no late payment* and *excellent credit rating*) =
0.80 + 0.45 – 0.40 = 0.85

**b.** For ARMs, we have the following information:

| | Late payment | No late payment | Total |
|---|---|---|---|
| **Excellent credit rating** | 0.45 | 0.05 | 0.50 |
| **Not excellent credit rating** | 0.20 | 0.30 | 0.50 |
| **Total** | 0.65 | 0.35 | 1.00 |

Using the addition law, we have:

P(*no late payment* or *excellent credit rating*) =
P(*no late payment*) + P(*excellent credit rating*) - P(*no late payment* and *excellent credit rating*) =
0.35 + 0.50 – 0.05 = 0.80

**E73.** **a. i.** $P(Republican) = 16/42 \approx 0.381$

**ii.** $P(voted\ Republican) = 13/42 \approx 0.310$

**iii.** $P(voted\ Republican \text{ and } Republican) = 12/42 \approx 0.286$

**iv.** $P(voted\ Republican \text{ or } Republican) = 17/42 \approx 0.405$

**v.** $P(Republican \mid voted\ Republican) = 12/13 \approx 0.923$

**vi.** $P(voted\ Republican \mid Republican) = 12/16 \approx 0.75$

**b.** No. P(*Republican*) = 0.381. P(*Republican* | *Voted Republican*) = 0.923. These are not equal.
**c.** No. There was one Republican who voted for a Democrat.

**E74.** **a.** $p + q$

**b.** $\frac{q}{p+q}$

**c.** $\frac{p}{p+r}$

**d.** $q$

**e.** $p + q + s$ (or $1 - r$)

**E75.** **a.**

| | *A* | *Not A* | Total |
|---|---|---|---|
| *B* | $\frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$ | $\frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16}$ | $\frac{5}{16}$ |
| *Not B* | $\frac{1}{8}$ | $\frac{9}{16}$ | $\frac{11}{16}$ |
| Total | $\frac{1}{4}$ | $\frac{3}{4}$ | 1 |

**b.** No. $P(B) = 5/16 \neq P(B \mid A)$

**E76.** The event *60 or more correct identifications of tap water* would occur only about 2.5% of the time if all of the tasters were guessing. Because this is so unlikely, it is the conclusion of Downhill Research that the tasters can distinguish between tap water and bottled water.

**E77.** The table summarizes the results of the study:

| | | Answer to Single Question | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Clinically Depressed? | Yes | 37 | 6 | 43 |
| | No | 8 | 28 | 36 |
| | Total | 45 | 34 | 79 |

The sensitivity of the single question is $\frac{37}{43}$, or about 86%. This means that 86% of the depressed patients will be identified as depressed by the single question.

The specificity of the single question is $\frac{28}{36}$, or about 78%. This means that 78% of the people who aren't depressed will be identified as not depressed by this question.

The positive predictive value is $\frac{37}{45}$, or about 82%. This means that 82% of the people identified as depressed by the single question actually are depressed.

The negative predictive value is $\frac{28}{34}$, or about 82%. This means that 82% of the people identified as not depressed by the single question are not depressed.

This is a reasonably good test, but the sensitivity is not as high as it should be. The number of false positives, 8, isn't much of a problem as these people will receive further testing. The number of false negatives, 6, is the bigger problem as these people won't receive the counseling they need. If depression is indeed a serious impediment to recovery from a stroke, a good screening test should have higher sensitivity than 86%.

**E78.** There are 8 choices for the first digit, and 9 choices for the second digit. The third digit must be different from the second, and can include 9 (which is not allowed for the second digit), so there are 9 choices for the third digit. So there are 8 · 9 · 9, or 648 possible area codes.

**E79. a.** There are 99,999 possible zip codes: 00001, 00002, . . . , 99999. The average number of people per zip code is 300,000,000 / 99,999 $\approx$ 3000.

**b.** There are 99,999 possible zip codes and 9,999 possible plus 4 codes. Thus, there are 99,999 · 9,999 = 999,890,001 possible zip plus 4 codes. The average number of people per zip plus 4 code is 300,000,000 / 999,890,001 $\approx$ 0.3. We each could have almost three zip plus 4 codes.

**E80.** In order to form a codon, there are three slots that must be filled with one of four nucleotides (A, G, C, or U). Since any of these values can be repeated in any slot, there are $4 \times 4 \times 4 = 64$ different codons.

## Concept Review Solutions

**C1.** C. It is far more likely that the teacher won't be abducted by extraterrestrials than that she will be abducted. It is only if we assume that the events are equally likely that we can conclude the probability is $\frac{1}{2}$ for each of the two events.

**C2.** E. P(1 on the first die) = P(1 on the first die | doubles) = $\frac{1}{6}$. The answer is not choice A, for example, because P(sum of 8) = $\frac{5}{36}$, but P(sum of 8 | doubles) = $\frac{1}{6}$

**C3.** B. It is impossible to get a sum of 3 and doubles on the same roll.

**C4.** C. Remember the choices are independent, so you multiply the probabilities. Moreover, the sample space size is decreased by 1 each time since we are making the choices without replacement; the number of outcomes corresponding to non-accidents also reduces by 1 with each such choice.

**C5.** E. This question cannot be answered without knowing if being female and getting an A are independent events. If all of the students getting A's were female, then the answer would be 30%; if none of the students getting A's were female, then the answer would be 0%. Any answer between 0% and 30% is possible.

**C6.** C This is easiest to see by constructing a table. The problem gives the percentages that go in the four cells. After filling in the cells, add across and down to get the marginal totals. The probability that a smoker gets lung cancer is $\frac{4}{26}$.

| | Smoker | Non-Smoker | Total |
|---|---|---|---|
| Lung Cancer | 4 | 8 | 12 |
| No Lung Cancer | 22 | 66 | 88 |
| Total | 26 | 74 | 100% |

**C7.** D. The easiest way to do this problem is by finding the complement:

$$1 - P(\text{none get it right}) = 1 - (0.3)(0.3)(0.3) = 0.973.$$

**C8.** D. There's about a 50% chance that the main antenna continues to function. When the main antenna fails (the remaining 50% of the time), there's a 20% chance of having a working backup antenna. The probability of at least one working antenna is then

$$0.5 + 0.5(0.2) = 0.6.$$

These probabilities also can be organized in a two-way table with one way being *survive/fail* for the main antenna and the other way being *survive/fail* for the backup antenna.

**C9.** Encourage students to tell an interesting story with only the calculations that are necessary. Students may make far more computations than they need to prove the important points, such as that although class mattered, being female was more important to survival. Some interesting probabilities to note are the following:

$$P(survived) = \frac{499}{1316} \approx 0.379$$

$$P(survived \mid 3rd\ class) = \frac{178}{706} \approx 0.252$$

$$P(survived \mid 1st\ class) = \frac{203}{325} \approx 0.625$$

$$P(survived \mid 1st\ class\ female) = \frac{141}{145} \approx 0.972$$

$$P(survived \mid 3rd\ class\ female) = \frac{90}{196} \approx 0.459$$

$$P(survived \mid 1st\ class\ male) = \frac{62}{180} \approx 0.344$$