

## Solutions to Chapter 1

1. The intent of this rather vague problem is to get you to compare the two notions, probability as intuition and relative frequency theory. There are many possible answers to how to make the statement "Ralph is probably guilty of theft" have a numerical value in the relative frequency theory. First step is to define a repeatable experiment along with its outcomes. The favorable outcome in this case would be 'guilty.' Repeating this experiment a large number of times would then give the desired probability in a relative frequency sense. We thus see that it may entail a lot of work to attach an objective numerical value to such a subjective statement, if in fact it can be done at all.

One possible approach would be to look through courthouse statistics for cases similar to Ralph's, similar both in terms of the case itself and the defendant. If we found a sufficiently large number of these cases, ten at least, we could then form the probability  $p = n_F/n$ , where  $n_F$  is the number of favorable (guilty) verdicts, and  $n$  is the total number of found cases. Here we effectively assume that the judge and jury are omniscient.

Another possibility is to find a large number of people with personalities and backgrounds similar to Ralph's, and to expose them to a very similar situation in which theft is possible. The fraction of these people that then steal in relation to the total number of people, would then give an objective meaning to the phrase "Ralph is probably guilty of theft."

2. Note that  $D \rightarrow 3$ , but  $3 \not\rightarrow D$ , i.e.,  $D$  implies 3 but not the other way around. Thus if we turn over card 2 and find a 3. So what? It was never stated that a  $3 \rightarrow F$ . Likewise, with card 3. On the other hand, if we turn over card 4 and find a  $D$ , then the rule is violated. Hence, we must turn over card 4 and card 1, of course.
3. First step here is to decide which kind of probability to use. Since no probabilities are explicitly given, it is reasonable to assume that all numbers are equally likely. Effectively we assume that the wheel is "fair." This then allows us to use the classical theory along with the axiomatic theory to solve this problem. Now we must find the corresponding probability model. We are told in the problem statement that the experiment is "spinning the wheel." We identify the pointed-to numbers as the outcomes  $\varsigma$ . The sample space is thus  $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ . The total number of outcomes is then 9. The probability of each elemental event  $\{i\}$  is then taken as  $P[\{i\}] \triangleq p = 1/9$ , as in the classical theory. We are also told in the problem statement that the contestant wins if an even number shows. The set of even numbers in  $\Omega$  is  $\{2, 4, 6, 8\}$ . We can write this event as a disjoint union of four singleton (atomic) events

$$\{2, 4, 6, 8\} = \{2\} \cup \{4\} \cup \{6\} \cup \{8\}.$$

Now we can apply axiom 3 of probability to write

$$\begin{aligned} P[\{2, 4, 6, 8\}] &= P[\{2\}] + P[\{4\}] + P[\{6\}] + P[\{8\}] \\ &= \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} \\ &= \frac{4}{9}. \end{aligned}$$

We have seen that some 'reasonable' assumptions are necessary to transform the given word problem into something that exactly corresponds to a probability model. It turns out that this is a general problem for such word problems, i.e. problems given in natural English.

4. The experiment involves flipping a fair coin 3 times. The outcome of each coin toss is either a head or a tail. Therefore, the sample space of the combined experiment that contains all the possible outcomes of the 3 tosses, is given by

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Since all the coins are fair, all the outcomes of the experiment are equally likely. The probability of each singleton event, i.e. an event with a single outcome, is then  $\frac{1}{8}$ . We are interested in finding the probability of the event  $A$ , which is the event of obtaining 2 heads and 1 tail. There are 3 favorable outcomes for this event given by  $A = \{HHT, HTH, THH\}$ . Therefore,  $P[A] = P[\{HHT\} \cup \{HTH\} \cup \{THH\}] = P[HHT] + P[HTH] + P[THH] = \frac{3}{8}$ . Note that we are able to write the probability of the event  $A$  as the sum of probability of the singleton events (from Axiom 3) because the singleton events of any experiment are mutually exclusive. Why?

5. The experiment contains drawing two balls (with replacement) from an urn containing balls numbered 1, 2, and 3. The sample space of the experiment is given by

$$\Omega = \{11, 12, 13, 21, 22, 23, 31, 32, 33\}.$$

The event of drawing a ball twice is said to occur when one of the outcomes 11, 22, or 33 occurs. Therefore, the event of drawing 2 equal balls  $E$ , is given by  $E = \{11, 22, 33\}$  and  $P[E] = P[\{11\}] + P[\{22\}] + P[\{33\}]$ . Since the balls are drawn at random, it can be assumed that drawing each ball is equally likely. Therefore, the singleton events, or equivalently outcomes of the experiment, are equally likely. Hence,  $P[E] = 3(\frac{1}{9}) = \frac{1}{3}$ .

6. Let  $b_1, b_2, \dots, b_6$  represent the six balls. Each outcome will be represented by the two balls that were drawn. In the first experiment, the balls are drawn without replacement; hence, the two balls drawn cannot have the same index. Then the sample space containing all the outcomes is given by

$$\begin{aligned} \Omega_1 = \{ & b_1b_2, b_1b_3, b_1b_4, b_1b_5, b_1b_6, \\ & b_2b_1, b_2b_3, b_2b_4, b_2b_5, b_2b_6, \\ & b_3b_1, b_3b_2, b_3b_4, b_3b_5, b_3b_6, \\ & b_4b_1, b_4b_2, b_4b_3, b_4b_5, b_4b_6, \\ & b_5b_1, b_5b_2, b_5b_3, b_5b_4, b_5b_6, \\ & b_6b_1, b_6b_2, b_6b_3, b_6b_4, b_6b_5\}. \end{aligned}$$

This can be written compactly as

$$\Omega_1 = \{b_{(i,j)} | 1 \leq i \leq 6, 1 \leq j \leq 6, i \neq j\}.$$

If the first ball is replaced before the second draw, then in addition to the outcomes in the earlier part, there are outcomes where both the two balls drawn are the same. The sample space for the new experiment is given by

$$\Omega_2 = \Omega_1 \cup \{b_1b_1, b_2b_2, b_3b_3, b_4b_4, b_5b_5, b_6b_6\}.$$

This can also be written as  $\Omega_2 = \{b_{(i,j)} | 1 \leq i \leq 6, 1 \leq j \leq 6\}$ .

7. Let  $h_M$  be the height of the man and  $h_W$  be the height of the woman. Each outcome of the experiment can be expressed as a two-tuple  $(h_M, h_W)$ . Thus

(a) The sample space  $\Omega$  is the set of all possible pairs of heights for the man and woman. This is given as

$$\Omega = \{(h_M, h_W) : h_M > 0, h_W > 0\}.$$

(b) The event  $E$ , which is a subset of  $\Omega$  is given by

$$E = \{(h_M, h_W) : h_M > 0, h_W > 0, h_M < h_W\}.$$

8. The word problem describes the physical experiment of drawing numbered balls from an urn. We need to find a corresponding mathematical model. First we form an appropriate event space with meaningful outcomes. Here the physical experiment is 'draw ball from urn,' so the outcome in words is 'particular labeled ball drawn,' which we can identify with its label. So we select as *outcome* in our mathematical model, the number on the drawn ball's face, i.e. the particular label. The outcomes are thus the integers 1,2,3,4,5,6,7,8, 9, and 10. The *sample space* is then  $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ , and is the set of all ten outcomes. We are told that  $E$  is 'the event of drawing a ball numbered no greater than 5.' Thus we define in our event field  $E = \{1, 2, 3, 4, 5\}$ . The other event specified in the word problem is  $F$  'the event of drawing a ball greater than 3 but less than 9.' In our mathematical event field this corresponds to  $F = \{4, 5, 6, 7, 8\}$ . Having constructed our sample space with indicated events, we can use elementary set theory to determine the following answers:

$$\begin{aligned} E^c &= \{6, 7, 8, 9, 10\}, & F^c &= \{1, 2, 3, 9, 10\}, \\ EF &= \{4, 5\}, & E \cup F &= \{1, 2, 3, 4, 5, 6, 7, 8\}, \\ EF^c &= \{1, 2, 3\}, & E^c F &= \{6, 7, 8\}, \\ E^c \cup F^c &= \{1, 2, 3, 6, 7, 8, 9, 10\}, \\ (EF^c) \cup (E^c F) &= \{1, 2, 3, 6, 7, 8\}, & (EF) \cup (E^c F^c) &= \{4, 5, 9, 10\} \\ (E \cup F)^c &= \{9, 10\}, & (EF)^c &= \{1, 2, 3, 6, 7, 8, 9, 10\}. \end{aligned}$$

The last part of the problem asks us to 'express these events in words.' Since we have a mathematical model, we should really more precisely ask what each of these events *corresponds to in words*. We know of course that  $E$  corresponds to 'drawing a ball numbered no greater than 5.' We can thus loosely write  $E = \{\text{'drawing a ball numbered no greater than 5'}\}$ , although in our mathematical model  $E$  is just the set of integers  $\{1, 2, 3, 4, 5\}$ . So when we write  $E = \{\text{'drawing a ball numbered no greater than 5'}\}$ , what we really mean is that the event  $E$  in our mathematical model corresponds to the physical event 'drawing a ball numbered no greater than 5' mentioned in the word problem. With this caveat in mind, we can then write:

$$\begin{aligned} E^c &= \{\text{'drawing a ball greater than 5'}\}, \\ F^c &= \{\text{'drawing a ball not in the range 4-8 inclusive'}\}, \\ EF &= \{\text{'drawing a ball greater than 3 and no greater than 5'}\}, \\ &\text{etc.} \end{aligned}$$

9. The sample space containing four equally likely outcomes is given by  $\Omega = \{\zeta_1, \zeta_2, \zeta_3, \zeta_4\}$ . Two events  $A = \{\zeta_1, \zeta_2\}$  and  $B = \{\zeta_2, \zeta_3\}$  are given. The required events can be easily obtained by observation.

$AB^c$  = set of outcomes in  $A$  and not in  $B = \{\zeta_1\}$ .

$BA^c$  = set of outcomes in  $B$  and not in  $A = \{\zeta_3\}$ .

$AB$  = set of outcomes in  $A$  and  $B = \{\zeta_2\}$ .

$A \cup B$  = set of outcomes in  $A$  or in  $B = \{\zeta_1, \zeta_2, \zeta_3\}$ .

10.  $A = AB \cup AB^c$ . This can be proved using the distributive law on

$$A = A\Omega = A(B \cup B^c) = AB \cup AB^c.$$

$A \cup B = (AB^c) \cup (BA^c) \cup (AB)$ . Here we first write  $A = A(B \cup B^c)$  and  $B = B(A \cup A^c)$ . Then we can write

$$\begin{aligned} A \cup B &= (A(B \cup B^c)) \cup (B(A \cup A^c)) \\ &= (AB \cup AB^c) \cup (BA \cup BA^c) \\ &= AB \cup AB^c \cup BA \cup BA^c \\ &= AB \cup AB^c \cup BA^c, \end{aligned}$$

using the above laws and formulas. Notice that the above two decompositions are into disjoint sets. From the third axiom of probability, we know that the probability of union of disjoint sets is the sum of the probabilities of the disjoint sets. Therefore, we can add the probabilities over the unions.

11. In a given random experiment there are four *equally likely* outcomes  $\zeta_1, \zeta_2, \zeta_3$ , and  $\zeta_4$ . Let the event  $A \triangleq \{\zeta_1, \zeta_2\}$ .  
 $P[A] = P[\{\zeta_1, \zeta_2\}] = P[\{\zeta_1\}] + P[\{\zeta_2\}] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ .  $A^c = \{\zeta_3, \zeta_4\}$ ,  
 $P[A^c] = P[\{\zeta_3, \zeta_4\}] = P[\{\zeta_3\}] + P[\{\zeta_4\}] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ .  
 Note that we are told that the four outcomes are equally likely. This means that the four singleton (atomic) events have equal probability.  $P[A] = \frac{1}{2} = 1 - P[A^c] = 1 - \frac{1}{2}$ .

12. (a) The three axioms of probability are given below

(a) [label=()]

(b) For any event  $A$ , the probability of the even occuring is always non-negative.

$$P[A] \geq 0.$$

This ensures that probability is never negative.

(c) The probability of occurrence of the sample space event  $\Omega$  is one.

$$P[\Omega] = 1.$$

This ensures that probability of no event exceeds one. The first two axioms ensures that the probability is a quantity between 0 and 1, inclusive.

(d) For any two events  $A, B$  that are disjoint, the probability of the union of the events is the sum of the probabilities of the two events.

$$P[A \cup B] = P[A] + P[B], \text{ when } AB = \phi.$$

This axiom tells us that the probability of any event can be obtained by the sum disjoint events that constitute the event.

(b) The event  $A \cup B$  can be obtained as the disjoint union of the three sets  $AB, AB^c, A^cB$ . Hence by applying the third axiom of probability, we obtain

$$\begin{aligned} P[A \cup B] &= P[AB \cup (AB^c \cup A^cB)] \\ &= P[AB] + P[AB^c \cup A^cB] \\ &= P[AB] + P[AB^c] + P[A^cB]. \end{aligned}$$

Now the event  $A$  can be written as the disjoint union of  $AB$  and  $AB^c$  (Axiom 3). Therefore

$$P[A] = P[AB] + P[AB^c] \implies P[AB^c] = P[A] - P[AB]$$

Similarly

$$P[B] = P[AB] + P[A^cB] \implies P[A^cB] = P[B] - P[AB].$$

Therefore  $P[A \cup B] = P[AB] + (P[A] - P[AB]) + (P[B] - P[AB]) = P[A] + P[B] - P[AB]$ .

13. We first form our mathematical model by setting outcomes  $\zeta = (\zeta_1, \zeta_2)$ , where  $\zeta_1$  corresponds to the label on the first ball drawn, and  $\zeta_2$  corresponds to the label on the second ball drawn. We can also write the outcomes as strings  $\zeta = \zeta_1\zeta_2$ . The sample space  $\Omega$  can then be identified with the 2-D array

11	12	13	14	15
21	22	23	24	25
31	32	33	34	35
41	42	43	44	45
51	52	53	54	55

There are thus 25 outcomes in the sample space. Now the word problem statement uses the phrase 'at random' to describe the drawing. This is a technical term that can be read 'equally likely.' Thus all the elementary events  $\{\zeta_1\zeta_2\}$  in our mathematical model must have equal probability, i.e.  $P[\{\zeta_1\zeta_2\}] = 1/25$ . Armed thusly we can attack the given problem as follows. Define the event  $E = \{\text{'sum of labels equals five'}\}$ , or precisely  $E = \{41, 32, 23, 14\}$ . Then we decompose this event into four singleton events as

$$E = \{41\} \cup \{32\} \cup \{23\} \cup \{14\}.$$

Since different singleton events are disjoint, probability adds, and we have

$$\begin{aligned} P[E] &= \frac{1}{25} + \frac{1}{25} + \frac{1}{25} + \frac{1}{25} \\ &= \frac{4}{25}. \end{aligned}$$

"Dim" ignored that outcome  $ij$  is different (distinguishable) from outcome  $ji$ . "Dense" talked about the sums and correctly noted that there were nine of them. However, he incorrectly assumed that each sum was equally likely. Looking at our sample space above, we can see that the sum 2 has only one favorable outcome 11, while the sum 6 has five favorable outcomes, just looking at the anti-diagonals of this matrix.

14. First we show  $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$ .

Let  $x \in A \cap (B \cup C)$ .

Then  $x \in A$  and  $x \in (B \cup C)$ .

$x \in A$  and  $x \in B$  or  $x \in C$ .

Say if  $x \in B$ . Then  $x \in A$  and  $x \in B$  (Step k)

Thus  $x \in (A \cap B)$ .

And therefore  $x \in (A \cap B) \cup (A \cap C)$ . Similar arguments can be made if we consider  $x \in C$  in step k, in which case we will show that  $x \in (A \cap C)$  and hence  $x \in (A \cap B) \cup (A \cap C)$ .

Thus we have shown that  $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$ .

Now we show that  $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$ .

Suppose  $x \in (A \cap B) \cup (A \cap C)$ . Then  $x \in (A \cap B)$  or  $x \in (A \cap C)$ .

Say  $x \in (A \cap B)$

Then  $x \in A$  and  $x \in B$ .

Or  $x \in A$  and  $x \in (B \cup C)$ . Or in other words,  $x \in A \cap (B \cup C)$ .

Similar arguments can be used to show that if  $x \in (A \cap C)$ , then  $x \in A \cap (B \cup C)$ .

Thus  $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$ .

Thus we have shown that both sets are contained in each other. Hence  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ .

15. We use the set identity  $\Omega = A \cup A^c$ . Since this union is disjoint, by the additivity of probability (i.e. axiom 3), we get  $1 = P[\Omega] = P[A] + P[A^c]$ , which with rearranging becomes the desired result.
16. (a)  $A \cap C = \{1, 2\} \cap \{4, 5, 6\} = \phi$ . Therefore,

$$\begin{aligned} P[A \cap C] &= P[\phi] \\ &= 1 - P[\Omega] \\ &\quad (\because \Omega \cap \Phi = \phi, 1 = P[\Omega \cup \phi] = P[\Omega] + P[\phi]) \\ &= 1 - 1 \quad (\text{because } P[\Omega] = 1) \\ &= 0. \end{aligned}$$

(b)  $P[A \cup B \cup C] = P[\{1, 2\} \cup \{2, 3\} \cup \{4, 5, 6\}] = P[\{1, 2, 3, 4, 5, 6\}] = P[\Omega] = 1$ .

(c) We see that  $B \cap C = \phi$  and so  $P[BC] = 0$ . For  $B$  and  $C$  to be independent,  $P[BC] = P[B]P[C]$ . Therefore, if either  $P[B] = 0$  or  $P[C] = 0$  or both are zeros,  $B$  and  $C$  will be independent.

17. This problem uses only set theory and just two axioms of probability to get these general results.

(a) We need to show  $P[\phi] = 0$ . We write the disjoint decomposition  $\Omega = \Omega \cup \phi$  and then use the additivity of probability (axiom 3) to get

$$\begin{aligned} P[\Omega] &= P[\Omega \cup \phi] \\ &= P[\Omega] + P[\phi]. \end{aligned}$$

So we must have  $P[\phi] = 0$ .

(b) Using set theory, we can write the disjoint decomposition

$$E = EF^c \cup EF.$$

Then by axiom 3, the additivity of probability, we have

$$\begin{aligned} P[E] &= P[EF^c \cup EF] \\ &= P[EF^c] + P[EF], \end{aligned}$$

or what is the same  $P[EF^c] = P[E] - P[EF]$ .

(c) Here we simply note  $E \cup E^c = \Omega$  is a disjoint decomposition, so that again by axiom 3,

$$\begin{aligned} P[\Omega] &= P[E] + P[E^c] \\ &= 1, \quad \text{by axiom 2,} \end{aligned}$$

which is the same as  $P[E] = 1 - P[E^c]$ .

18. The outcome is the result of a probabilistic experiment. An event is a collection (set) of outcomes. The field of events is the complete collection of events that are relevant for the given probability problem.

19. We start with the mutually exclusive decomposition

$$A \cup B = AB^c \cup AB \cup A^cB,$$

yielding  $P[A \cup B] = P[AB^c] + P[AB] + P[A^cB]$ . Then consider the two simple disjoint decompositions

$$A = AB^c \cup AB \quad \text{and} \quad B = A^cB \cup AB,$$

which yield  $P[A] = P[AB^c] + P[AB]$  and  $P[B] = P[A^cB] + P[AB]$ . Putting them all together, we have

$$\begin{aligned} P[A \cup B] &= P[AB^c] + P[AB] + P[A^cB] \\ &= (P[A] - P[AB]) + P[AB] + (P[B] - P[AB]) \\ &= P[A] + P[B] - P[AB]. \end{aligned}$$

20. From Eq. 1.4-3, we see that  $E \oplus F = (E - F) \cup (F - E) = EF^c \cup E^cF$ . We see that  $EF^c$  and  $E^cF$  are disjoint, i.e.,  $(EF^c) \cap (E^cF) = \phi$ . Therefore, the probability of the union of  $EF^c$  and  $E^cF$  are the sum of the probabilities of the two events. In other words,

$$P(E \oplus F) = P(EF^c \cup E^cF) = P(EF^c) + P(E^cF).$$

21. We have already (Problem 17) seen that we can write  $P[EF^c] = P[E] - P[EF]$  and  $P[E^cF] = P[F] - P[EF]$ . Therefore,  $P(E \oplus F) = P(EF^c) + P(E^cF) = P[E] + P[F] - 2P[EF]$ .

22. (a) For simplicity associate as follows: cat=1, dog=2, goat=3, and pig=4. The outcomes  $\xi$  then become the integers 1,2,3, and 4. The sample space  $\Omega = \{1, 2, 3, 4\}$ . For probability information we are given:

$$P[\{1, 2\}] = 0.9, P[\{3, 4\}] = 0.1, P[\{4\}] = 0.05, \text{ and } P[\{2\}] = 0.5.$$

Now for every event in our field of events, we must be able to specify the probability. This is equivalent to being able to supply the probability for all the singleton events. To see if we can do this, we note that singleton events  $\{1\}$  and  $\{3\}$  are missing probabilities, so we first write

$$\begin{aligned} \{1\} &= \{1, 2\} - \{2\}, \text{ so that} \\ P[\{1\}] &= P[\{1, 2\}] - P[\{2\}] = 0.9 - 0.5 = 0.4. \end{aligned}$$