PROBLEMS

SECTION 1.1. Sets

Problem 1. Consider rolling a six-sided die. Let A be the set of outcomes where the roll is an even number. Let B be the set of outcomes where the roll is greater than 3. Calculate and compare the sets on both sides of De Morgan's laws

$$(A \cup B)^c = A^c \cap B^c, \qquad \big(A \cap B\big)^c = A^c \cup B^c.$$

Problem 2. Let A and B be two sets.

(a) Show that

$$A^c = (A^c \cap B) \cup (A^c \cap B^c), \qquad B^c = (A \cap B^c) \cup (A^c \cap B^c).$$

(b) Show that

$$(A \cap B)^c = (A^c \cap B) \cup (A^c \cap B^c) \cup (A \cap B^c).$$

(c) Consider rolling a fair six-sided die. Let A be the set of outcomes where the roll is an odd number. Let B be the set of outcomes where the roll is less than 4. Calculate the sets on both sides of the equality in part (b), and verify that the equality holds.

Problem 3.* Prove the identity

$$A \cup \big(\cap_{n=1}^{\infty} B_n\big) = \cap_{n=1}^{\infty}(A \cup B_n).$$

Solution. If x belongs to the set on the left, there are two possibilities. Either $x \in A$, in which case x belongs to all of the sets $A \cup B_n$, and therefore belongs to the set on the right. Alternatively, x belongs to all of the sets B_n in which case, it belongs to all of the sets $A \cup B_n$, and therefore again belongs to the set on the right.

Conversely, if x belongs to the set on the right, then it belongs to $A \cup B_n$ for all n. If x belongs to A, then it belongs to the set on the left. Otherwise, x must belong to every set B_n and again belongs to the set on the left.

Problem 4.* Cantor's diagonalization argument. Show that the unit interval $[0, 1]$ is uncountable, i.e., its elements cannot be arranged in a sequence.

Solution. Any number x in $[0, 1]$ can be represented in terms of its decimal expansion, e.g., $1/3 = 0.3333\cdots$. Note that most numbers have a unique decimal expansion, but there are a few exceptions. For example, $1/2$ can be represented as $0.5000\cdots$ or as $0.49999\cdots$. It can be shown that this is the only kind of exception, i.e., decimal expansions that end with an infinite string of zeroes or an infinite string of nines.

Suppose, to obtain a contradiction, that the elements of $[0,1]$ can be arranged in a sequence $x_1, x_2, x_3, \ldots$, so that every element of $[0,1]$ appears in the sequence. Consider the decimal expansion of x_n:

$$x_n = 0.a_n^1 a_n^2 a_n^3 \cdots,$$

where each digit a_n^i belongs to $\{0, 1, \ldots, 9\}$. Consider now a number y constructed as follows. The nth digit of y can be 1 or 2, and is chosen so that it is different from the nth digit of x_n. Note that y has a unique decimal expansion since it does not end with an infinite sequence of zeroes or nines. The number y differs from each x_n, since it has a different nth digit. Therefore, the sequence $x_1, x_2, \ldots$ does not exhaust the elements of $[0,1]$, contrary to what was assumed. The contradiction establishes that the set $[0,1]$ is uncountable.

SECTION 1.2. Probabilistic Models

Problem 5. Out of the students in a class, 60% are geniuses, 70% love chocolate, and 40% fall into both categories. Determine the probability that a randomly selected student is neither a genius nor a chocolate lover.

Problem 6. A six-sided die is loaded in a way that each even face is twice as likely as each odd face. All even faces are equally likely, as are all odd faces. Construct a probabilistic model for a single roll of this die and find the probability that the outcome is less than 4.

Problem 7. A four-sided die is rolled repeatedly, until the first time (if ever) that an even number is obtained. What is the sample space for this experiment?

Problem 8. You enter a special kind of chess tournament, in which you play one game with each of three opponents, but you get to choose the order in which you play your opponents, knowing the probability of a win against each. You win the tournament if you win two games in a row, and you want to maximize the probability of winning. Show that it is optimal to play the weakest opponent second, and that the order of playing the other two opponents does not matter.

Problem 9. A partition of the sample space Ω is a collection of disjoint events $S_1, \ldots, S_n$ such that $\Omega = \cup_{i=1}^n S_i$.

(a) Show that for any event A, we have

$$\mathbf{P}(A) = \sum_{i=1}^{n} \mathbf{P}(A \cap S_i).$$

(b) Use part (a) to show that for any events A, B, and C, we have

$$\mathbf{P}(A) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap C) + \mathbf{P}\big(A \cap B^c \cap C^c\big) - \mathbf{P}(A \cap B \cap C).$$

Problem 10. Show the formula

$$\mathbf{P}\big((A \cap B^c) \cup (A^c \cap B)\big) = \mathbf{P}(A) + \mathbf{P}(B) - 2\mathbf{P}(A \cap B),$$

which gives the probability that exactly one of the events A and B will occur. [Compare with the formula $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$, which gives the probability that at least one of the events A and B will occur.]

Problem 11.* Bonferroni's inequality.

(a) Prove that for any two events A and B, we have

$$\mathbf{P}(A \cap B) \geq \mathbf{P}(A) + \mathbf{P}(B) - 1.$$

(b) Generalize to the case of n events $A_1, A_2, \ldots, A_n$, by showing that

$$\mathbf{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \geq \mathbf{P}(A_1) + \mathbf{P}(A_2) + \cdots + \mathbf{P}(A_n) - (n-1).$$

Solution. We have $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$ and $\mathbf{P}(A \cup B) \leq 1$, which implies part (a). For part (b), we use De Morgan's law to obtain

$$\begin{aligned} 1 - \mathbf{P}(A_1 \cap \cdots \cap A_n) &= \mathbf{P}\big((A_1 \cap \cdots \cap A_n)^c\big) \\ &= \mathbf{P}(A_1^c \cup \cdots \cup A_n^c) \\ &\leq \mathbf{P}(A_1^c) + \cdots + \mathbf{P}(A_n^c) \\ &= \big(1 - \mathbf{P}(A_1)\big) + \cdots + \big(1 - \mathbf{P}(A_n)\big) \\ &= n - \mathbf{P}(A_1) - \cdots - \mathbf{P}(A_n). \end{aligned}$$

Problem 12.* The inclusion-exclusion formula. Show the following generalizations of the formula

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B).$$

(a) Let A, B, and C be events. Then,

$$\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(A \cap B) - \mathbf{P}(B \cap C) - \mathbf{P}(A \cap C) + \mathbf{P}(A \cap B \cap C).$$

(b) Let $A_1, A_2, \ldots, A_n$ be events. Let $S_1 = \{i \mid 1 \leq i \leq n\}$, $S_2 = \{(i_1, i_2) \mid 1 \leq i_1 < i_2 \leq n\}$, and more generally, let S_m be the set of all m-tuples $(i_1, \ldots, i_m)$ of indices that satisfy $1 \leq i_1 < i_2 < \cdots < i_m \leq n$. Then,

$$\begin{aligned} \mathbf{P}\left(\cup_{k=1}^n A_k\right) = \sum_{i \in S_1} \mathbf{P}(A_i) &- \sum_{(i_1, i_2) \in S_2} \mathbf{P}(A_{i_1} \cap A_{i_2}) \\ &+ \sum_{(i_1, i_2, i_3) \in S_3} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \cdots + (-1)^{n-1} \mathbf{P}\left(\cap_{k=1}^n A_k\right). \end{aligned}$$

Solution. (a) We use the formulas $\mathbf{P}(X \cup Y) = \mathbf{P}(X) + \mathbf{P}(Y) - \mathbf{P}(X \cap Y)$ and $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$. We have

$$\begin{aligned} \mathbf{P}(A \cup B \cup C) &= \mathbf{P}(A \cup B) + \mathbf{P}(C) - \mathbf{P}\big((A \cup B) \cap C\big) \\ &= \mathbf{P}(A \cup B) + \mathbf{P}(C) - \mathbf{P}\big((A \cap C) \cup (B \cap C)\big) \\ &= \mathbf{P}(A \cup B) + \mathbf{P}(C) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) + \mathbf{P}(A \cap B \cap C) \\ &= \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) + \mathbf{P}(C) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) \\ &\quad + \mathbf{P}(A \cap B \cap C) \\ &= \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(A \cap B) - \mathbf{P}(B \cap C) - \mathbf{P}(A \cap C) \\ &\quad + \mathbf{P}(A \cap B \cap C). \end{aligned}$$

(b) Use induction and verify the main induction step by emulating the derivation of part (a). For a different approach, see the problems at the end of Chapter 2.

Problem 13.* **Continuity property of probabilities.**

(a) Let $A_1, A_2, \ldots$ be an infinite sequence of events, which is "monotonically increasing," meaning that $A_n \subset A_{n+1}$ for every n. Let $A = \cup_{n=1}^{\infty} A_n$. Show that $\mathbf{P}(A) = \lim_{n\to\infty} \mathbf{P}(A_n)$. *Hint:* Express the event A as a union of countably many disjoint sets.

(b) Suppose now that the events are "monotonically decreasing," i.e., $A_{n+1} \subset A_n$ for every n. Let $A = \cap_{n=1}^{\infty} A_n$. Show that $\mathbf{P}(A) = \lim_{n\to\infty} \mathbf{P}(A_n)$. *Hint:* Apply the result of part (a) to the complements of the events.

(c) Consider a probabilistic model whose sample space is the real line. Show that

$$\mathbf{P}\big([0,\infty)\big) = \lim_{n\to\infty} \mathbf{P}\big([0,n]\big), \qquad \text{and} \qquad \lim_{n\to\infty} \mathbf{P}\big([n,\infty)\big) = 0.$$

Solution. (a) Let $B_1 = A_1$ and, for $n \geq 2$, $B_n = A_n \cap A_{n-1}^c$. The events B_n are disjoint, and we have $\cup_{k=1}^{n} B_k = A_n$, and $\cup_{k=1}^{\infty} B_k = A$. We apply the additivity axiom to obtain

$$\mathbf{P}(A) = \sum_{k=1}^{\infty} \mathbf{P}(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{P}(B_k) = \lim_{n\to\infty} \mathbf{P}(\cup_{k=1}^{n} B_k) = \lim_{n\to\infty} \mathbf{P}(A_n).$$

(b) Let $C_n = A_n^c$ and $C = A^c$. Since $A_{n+1} \subset A_n$, we obtain $C_n \subset C_{n+1}$, and the events C_n are increasing. Furthermore, $C = A^c = (\cap_{n=1}^{\infty} A_n)^c = \cup_{n=1}^{\infty} A_n^c = \cup_{n=1}^{\infty} C_n$. Using the result from part (a) for the sequence C_n, we obtain

$$1 - \mathbf{P}(A) = \mathbf{P}(A^c) = \mathbf{P}(C) = \lim_{n\to\infty} \mathbf{P}(C_n) = \lim_{n\to\infty} \big(1 - \mathbf{P}(A_n)\big),$$

from which we conclude that $\mathbf{P}(A) = \lim_{n\to\infty} \mathbf{P}(A_n)$.

(c) For the first equality, use the result from part (a) with $A_n = [0, n]$ and $A = [0, \infty)$. For the second, use the result from part (b) with $A_n = [n, \infty)$ and $A = \cap_{n=1}^{\infty} A_n = \emptyset$.

SECTION 1.3. Conditional Probability

Problem 14. We roll two fair 6-sided dice. Each one of the 36 possible outcomes is assumed to be equally likely.

(a) Find the probability that doubles are rolled.

(b) Given that the roll results in a sum of 4 or less, find the conditional probability that doubles are rolled.

(c) Find the probability that at least one die roll is a 6.

(d) Given that the two dice land on different numbers, find the conditional probability that at least one die roll is a 6.

Problem 15. A coin is tossed twice. Alice claims that the event of two heads is at least as likely if we know that the first toss is a head than if we know that at least one of the tosses is a head. Is she right? Does it make a difference if the coin is fair or unfair? How can we generalize Alice's reasoning?

Problem 16. We are given three coins: one has heads in both faces, the second has tails in both faces, and the third has a head in one face and a tail in the other. We choose a coin at random, toss it, and the result is heads. What is the probability that the opposite face is tails?

Problem 17. A batch of one hundred items is inspected by testing four randomly selected items. If one of the four is defective, the batch is rejected. What is the probability that the batch is accepted if it contains five defectives?

Problem 18. Let A and B be events. Show that $\mathbf{P}(A \cap B \mid B) = \mathbf{P}(A \mid B)$, assuming that $\mathbf{P}(B) > 0$.

SECTION 1.4. Total Probability Theorem and Bayes' Rule

Problem 19. Alice searches for her term paper in her filing cabinet. which has several drawers. She knows that she left her term paper in drawer j with probability $p_j > 0$. The drawers are so messy that even if she correctly guesses that the term paper is in drawer i. the probability that she finds it is only d_i. Alice searches in a particular drawer. say drawer i. but the search is unsuccessful. Conditioned on this event, show that the probability that her paper is in drawer j, is given by

$$\frac{p_j}{1 - p_i d_i}, \qquad \text{if } j \neq i, \qquad\qquad \frac{p_i(1 - d_i)}{1 - p_i d_i}, \qquad \text{if } j = i.$$

Problem 20. How an inferior player with a superior strategy can gain an advantage. Boris is about to play a two-game chess match with an opponent, and wants to find the strategy that maximizes his winning chances. Each game ends with either a win by one of the players, or a draw. If the score is tied at the end of the two games, the match goes into sudden-death mode, and the players continue to play until the first time one of them wins a game (and the match). Boris has two playing styles. *timid* and *bold*, and he can choose one of the two at will in each game. no matter what style he chose in previous games. With timid play. he draws with probability $p_d > 0$, and he loses with probability $1 - p_d$. With bold play. he wins with probability p_w, and he loses with probability $1 - p_w$. Boris will always play bold during sudden death, but may switch style between games 1 and 2.

(a) Find the probability that Boris wins the match for each of the following strategies:
   (i) Play bold in both games 1 and 2.
   (ii) Play timid in both games 1 and 2.
   (iii) Play timid whenever he is ahead in the score. and play bold otherwise.

(b) Assume that $p_w < 1/2$, so Boris is the worse player, regardless of the playing style he adopts. Show that with the strategy in (iii) above. and depending on the values of p_w and p_d. Boris may have a better than a 50-50 chance to win the match. How do you explain this advantage?

Problem 21. Two players take turns removing a ball from a jar that initially contains m white and n black balls. The first player to remove a white ball wins. Develop a recursive formula that allows the convenient computation of the probability that the starting player wins.

Problem 22. Each of k jars contains m white and n black balls. A ball is randomly chosen from jar 1 and transferred to jar 2, then a ball is randomly chosen from jar 2 and transferred to jar 3, etc. Finally, a ball is randomly chosen from jar k. Show that the probability that the last ball is white is the same as the probability that the first ball is white, i.e., it is $m/(m+n)$.

Problem 23. We have two jars, each initially containing an equal number of balls. We perform four successive ball exchanges. In each exchange, we pick simultaneously and at random a ball from each jar and move it to the other jar. What is the probability that at the end of the four exchanges all the balls will be in the jar where they started?

Problem 24. **The prisoner's dilemma.** The release of two out of three prisoners has been announced, but their identity is kept secret. One of the prisoners considers asking a friendly guard to tell him who is the prisoner other than himself that will be released, but hesitates based on the following rationale: at the prisoner's present state of knowledge, the probability of being released is 2/3, but after he knows the answer, the probability of being released will become 1/2, since there will be two prisoners (including himself) whose fate is unknown and exactly one of the two will be released. What is wrong with this line of reasoning?

Problem 25. **A two-envelopes puzzle.** You are handed two envelopes, and you know that each contains a positive integer dollar amount and that the two amounts are different. The values of these two amounts are modeled as constants that are unknown. Without knowing what the amounts are, you select at random one of the two envelopes, and after looking at the amount inside, you may switch envelopes if you wish. A friend claims that the following strategy will increase above 1/2 your probability of ending up with the envelope with the larger amount: toss a coin repeatedly, let X be equal to 1/2 plus the number of tosses required to obtain heads for the first time, and switch if the amount in the envelope you selected is less than the value of X. Is your friend correct?

Problem 26. **The paradox of induction.** Consider a statement whose truth is unknown. If we see many examples that are compatible with it, we are tempted to view the statement as more probable. Such reasoning is often referred to as *inductive inference* (in a philosophical, rather than mathematical sense). Consider now the statement that "all cows are white." An equivalent statement is that "everything that is not white is not a cow." We then observe several black crows. Our observations are clearly compatible with the statement, but do they make the hypothesis "all cows are white" more likely?

To analyze such a situation, we consider a probabilistic model. Let us assume that there are two possible states of the world, which we model as complementary events:

$$A : \text{all cows are white},$$

$$A^c : 50\% \text{ of all cows are white}.$$

Let p be the prior probability $\mathbf{P}(A)$ that all cows are white. We make an observation of a cow or a crow, with probability q and $1-q$, respectively, independent of whether event A occurs or not. Assume that $0 < p < 1$, $0 < q < 1$, and that all crows are black.

(a) Given the event $B = \{\text{a black crow was observed}\}$, what is $\mathbf{P}(A \mid B)$?

(b) Given the event $C = \{\text{a white cow was observed}\}$, what is $\mathbf{P}(A \mid C)$?

Problem 27. Alice and Bob have $2n + 1$ coins, each coin with probability of heads equal to $1/2$. Bob tosses $n+1$ coins, while Alice tosses the remaining n coins. Assuming independent coin tosses, show that the probability that after all coins have been tossed, Bob will have gotten more heads than Alice is $1/2$.

Problem 28.* Conditional version of the total probability theorem. Let $C_1, \ldots, C_n$ be disjoint events that form a partition of the state space. Let also A and B be events such that $\mathbf{P}(B \cap C_i) > 0$ for all i. Show that

$$\mathbf{P}(A \mid B) = \sum_{i=1}^{n} \mathbf{P}(C_i \mid B)\mathbf{P}(A \mid B \cap C_i).$$

Solution. We have

$$\mathbf{P}(A \cap B) = \sum_{i=1}^{n} \mathbf{P}\big((A \cap B) \cap C_i\big),$$

and by using the multiplication rule,

$$\mathbf{P}\big((A \cap B) \cap C_i\big) = \mathbf{P}(B)\mathbf{P}(C_i \mid B)\mathbf{P}(A \mid B \cap C_i).$$

Combining these two equations, dividing by $\mathbf{P}(B)$, and using the formula $\mathbf{P}(A \mid B) = \mathbf{P}(A \cap B)/\mathbf{P}(B)$, we obtain the desired result.

Problem 29.* Let A and B be events with $\mathbf{P}(A) > 0$ and $\mathbf{P}(B) > 0$. We say that an event B *suggests* an event A if $\mathbf{P}(A \mid B) > \mathbf{P}(A)$, and *does not suggest* event A if $\mathbf{P}(A \mid B) < \mathbf{P}(A)$.

(a) Show that B suggests A if and only if A suggests B.

(b) Assume that $\mathbf{P}(B^c) > 0$. Show that B suggests A if and only if B^c does not suggest A.

(c) We know that a treasure is located in one of two places, with probabilities β and $1 - \beta$, respectively, where $0 < \beta < 1$. We search the first place and if the treasure is there, we find it with probability $p > 0$. Show that the event of not finding the treasure in the first place suggests that the treasure is in the second place.

Solution. (a) We have $\mathbf{P}(A \mid B) = \mathbf{P}(A \cap B)/\mathbf{P}(B)$, so B suggests A if and only if $\mathbf{P}(A \cap B) > \mathbf{P}(A)\mathbf{P}(B)$, which is equivalent to A suggesting B, by symmetry.

(b) Since $\mathbf{P}(B) + \mathbf{P}(B^c) = 1$, we have

$$\mathbf{P}(B)\mathbf{P}(A) + \mathbf{P}(B^c)\mathbf{P}(A) = \mathbf{P}(A) = \mathbf{P}(B)\mathbf{P}(A \mid B) + \mathbf{P}(B^c)\mathbf{P}(A \mid B^c),$$

which implies that

$$\mathbf{P}(B^c)\big(\mathbf{P}(A) - \mathbf{P}(A \mid B^c)\big) = \mathbf{P}(B)\big(\mathbf{P}(A \mid B) - \mathbf{P}(A)\big).$$

Thus, $\mathbf{P}(A \mid B) > \mathbf{P}(A)$ (B suggests A) if and only if $\mathbf{P}(A) > \mathbf{P}(A \mid B^c)$ (B^c does not suggest A).

(c) Let A and B be the events

$$A = \{\text{the treasure is in the second place}\},$$
$$B = \{\text{we don't find the treasure in the first place}\}.$$

Using the total probability theorem, we have

$$\mathbf{P}(B) = \mathbf{P}(A^c)\mathbf{P}(B \mid A^c) + \mathbf{P}(A)\mathbf{P}(B \mid A) = \beta(1-p) + (1-\beta),$$

so

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{1-\beta}{\beta(1-p)+(1-\beta)} = \frac{1-\beta}{1-\beta p} > 1-\beta = \mathbf{P}(A).$$

It follows that event B suggests event A.

SECTION 1.5. Independence

Problem 30. A hunter has two hunting dogs. One day, on the trail of some animal, the hunter comes to a place where the road diverges into two paths. He knows that each dog, independent of the other, will choose the correct path with probability p. The hunter decides to let each dog choose a path, and if they agree, take that one, and if they disagree, to randomly pick a path. Is his strategy better than just letting one of the two dogs decide on a path?

Problem 31. Communication through a noisy channel. A source transmits a message (a string of symbols) through a noisy communication channel. Each symbol is 0 or 1 with probability p and $1-p$, respectively, and is received incorrectly with probability ϵ_0 and ϵ_1, respectively (see Fig. 1.18). Errors in different symbol transmissions are independent.

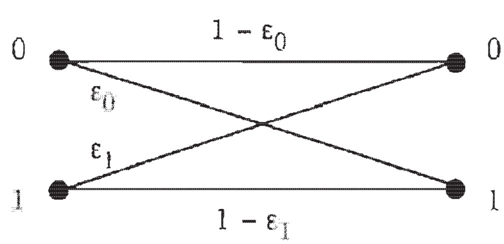

Figure 1.18: Error probabilities in a binary communication channel.

(a) What is the probability that the kth symbol is received correctly?

(b) What is the probability that the string of symbols 1011 is received correctly?

(c) In an effort to improve reliability, each symbol is transmitted three times and the received string is decoded by majority rule. In other words, a 0 (or 1) is transmitted as 000 (or 111, respectively), and it is decoded at the receiver as a 0 (or 1) if and only if the received three-symbol string contains at least two 0s (or 1s, respectively). What is the probability that a 0 is correctly decoded?

(d) For what values of ϵ_0 is there an improvement in the probability of correct decoding of a 0 when the scheme of part (c) is used?

(e) Suppose that the scheme of part (c) is used. What is the probability that a symbol was 0 given that the received string is 101?

Problem 32. The king's sibling. The king has only one sibling. What is the probability that the sibling is male? Assume that every birth results in a boy with probability 1/2, independent of other births. Be careful to state any additional assumptions you have to make in order to arrive at an answer.

Problem 33. Using a biased coin to make an unbiased decision. Alice and Bob want to choose between the opera and the movies by tossing a fair coin. Unfortunately, the only available coin is biased (though the bias is not known exactly). How can they use the biased coin to make a decision so that either option (opera or the movies) is equally likely to be chosen?

Problem 34. An electrical system consists of identical components, each of which is operational with probability p, independent of other components. The components are connected in three subsystems, as shown in Fig. 1.19. The system is operational if there is a path that starts at point A, ends at point B, and consists of operational components. What is the probability of this happening?

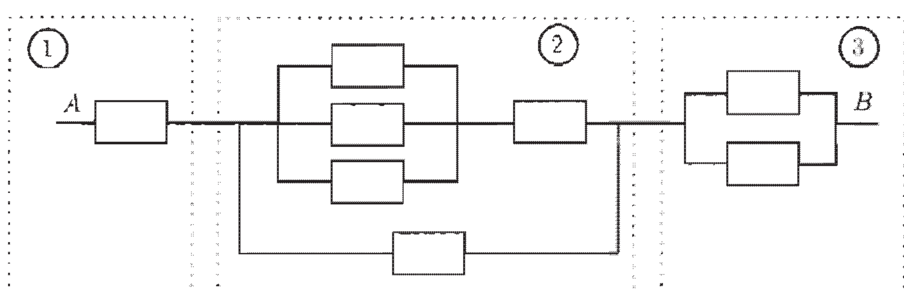

Figure 1.19: A system of identical components that consists of the three subsystems 1, 2, and 3. The system is operational if there is a path that starts at point A, ends at point B, and consists of operational components.

Problem 35. Reliability of a k-out-of-n system. A system consists of n identical components, each of which is operational with probability p, independent of other components. The system is operational if at least k out of the n components are operational. What is the probability that the system is operational?

Problem 36. A power utility can supply electricity to a city from n different power plants. Power plant i fails with probability p_i, independent of the others.

(a) Suppose that any one plant can produce enough electricity to supply the entire city. What is the probability that the city will experience a black-out?

(b) Suppose that two power plants are necessary to keep the city from a black-out. Find the probability that the city will experience a black-out.

Problem 37. A cellular phone system services a population of n_1 "voice users" (those who occasionally need a voice connection) and n_2 "data users" (those who occasionally need a data connection). We estimate that at a given time, each user will need to be connected to the system with probability p_1 (for voice users) or p_2 (for data users), independent of other users. The data rate for a voice user is r_1 bits/sec and for a data user is r_2 bits/sec. The cellular system has a total capacity of c bits/sec. What is the probability that more users want to use the system than the system can accommodate?

Problem 38. The problem of points. Telis and Wendy play a round of golf (18 holes) for a \$10 stake, and their probabilities of winning on any one hole are p and $1 - p$, respectively, independent of their results in other holes. At the end of 10 holes, with the score 4 to 6 in favor of Wendy, Telis receives an urgent call and has to report back to work. They decide to split the stake in proportion to their probabilities of winning had they completed the round, as follows. If p_T and p_W are the conditional probabilities that Telis and Wendy, respectively, are ahead in the score after 18 holes given the 4-6 score after 10 holes, then Telis should get a fraction $p_T/(p_T + p_W)$ of the stake, and Wendy should get the remaining $p_W/(p_T + p_W)$. How much money should Telis get? *Note*: This is an example of the, so-called, problem of points, which played an important historical role in the development of probability theory. The problem was posed by Chevalier de Méré in the 17th century to Pascal, who introduced the idea that the stake of an interrupted game should be divided in proportion to the players' conditional probabilities of winning given the state of the game at the time of interruption. Pascal worked out some special cases and through a correspondence with Fermat, stimulated much thinking and several probability-related investigations.

Problem 39. A particular class has had a history of low attendance. The annoyed professor decides that she will not lecture unless at least k of the n students enrolled in the class are present. Each student will independently show up with probability p_g if the weather is good, and with probability p_b if the weather is bad. Given the probability of bad weather on a given day, obtain an expression for the probability that the professor will teach her class on that day.

Problem 40. Consider a coin that comes up heads with probability p and tails with probability $1-p$. Let q_n be the probability that after n independent tosses, there have been an even number of heads. Derive a recursion that relates q_n to q_{n-1}, and solve this recursion to establish the formula

$$q_n = \big(1 + (1 - 2p)^n\big)/2.$$

Problem 41. Consider a game show with an infinite pool of contestants, where at each round i, contestant i obtains a number by spinning a continuously calibrated wheel. The contestant with the smallest number thus far survives. Successive wheel spins are independent and we assume that there are no ties. Let N be the round at which contestant 1 is eliminated. For any positive integer n, find $\mathbf{P}(N = n)$.

Problem 42.* Gambler's ruin. A gambler makes a sequence of independent bets. In each bet, he wins \$1 with probability p, and loses \$1 with probability $1-p$. Initially, the gambler has \$$k$, and plays until he either accumulates \$$n$ or has no money left. What is the probability that the gambler will end up with \$$n$?

Solution. Let us denote by A the event that he ends up with \$$n$, and by F the event that he wins the first bet. Denote also by w_k the probability of event A, if he starts with \$$k$. We apply the total probability theorem to obtain

$$w_k = \mathbf{P}(A \mid F)\mathbf{P}(F) + \mathbf{P}(A \mid F^c)\mathbf{P}(F^c) = p\mathbf{P}(A \mid F) + q\mathbf{P}(A \mid F^c). \qquad 0 < k < n,$$

where $q = 1-p$. By the independence of past and future bets, having won the first bet is the same as if he were just starting now but with \$$(k+1)$, so that $\mathbf{P}(A \mid F) = w_{k+1}$ and similarly $\mathbf{P}(A \mid F^c) = w_{k-1}$. Thus, we have $w_k = pw_{k+1} + qw_{k-1}$, which can be written as

$$w_{k+1} - w_k = r(w_k - w_{k-1}), \qquad 0 < k < n,$$

where $r = q/p$. We will solve for w_k in terms of p and q using iteration, and the boundary values $w_0 = 0$ and $w_n = 1$.

We have $w_{k+1} - w_k = r^k(w_1 - w_0)$, and since $w_0 = 0$,

$$w_{k+1} = w_k + r^k w_1 = w_{k-1} + r^{k-1} w_1 + r^k w_1 = w_1 + rw_1 + \cdots + r^k w_1.$$

The sum in the right-hand side can be calculated separately for the two cases where $r = 1$ (or $p = q$) and $r \neq 1$ (or $p \neq q$). We have

$$w_k = \begin{cases} \dfrac{1-r^k}{1-r} w_1, & \text{if } p \neq q. \\ kw_1. & \text{if } p = q. \end{cases}$$

Since $w_n = 1$, we can solve for w_1 and therefore for w_k:

$$w_1 = \begin{cases} \dfrac{1-r}{1-r^n}, & \text{if } p \neq q, \\ \dfrac{1}{n}, & \text{if } p = q, \end{cases}$$

so that

$$w_k = \begin{cases} \dfrac{1-r^k}{1-r^n}, & \text{if } p \neq q, \\ \dfrac{k}{n}, & \text{if } p = q. \end{cases}$$

Problem 43.* Let A and B be independent events. Use the definition of independence to prove the following:

(a) The events A and B^c are independent.

(b) The events A^c and B^c are independent.

Solution. (a) The event A is the union of the disjoint events $A \cap B^c$ and $A \cap B$. Using the additivity axiom and the independence of A and B, we obtain

$$\mathbf{P}(A) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap B^c) = \mathbf{P}(A)\mathbf{P}(B) + \mathbf{P}(A \cap B^c).$$

It follows that

$$\mathbf{P}(A \cap B^c) = \mathbf{P}(A)\big(1 - \mathbf{P}(B)\big) = \mathbf{P}(A)\mathbf{P}(B^c).$$

so A and B^c are independent.

(b) Apply the result of part (a) twice: first on A and B. then on B^c and A.

Problem 44.* Let A. B. and C be independent events, with $\mathbf{P}(C) > 0$. Prove that A and B are conditionally independent given C.

Solution. We have

$$\begin{aligned}\mathbf{P}(A \cap B \mid C) &= \frac{\mathbf{P}(A \cap B \cap C)}{\mathbf{P}(C)} \\ &= \frac{\mathbf{P}(A)\mathbf{P}(B)\mathbf{P}(C)}{\mathbf{P}(C)} \\ &= \mathbf{P}(A)\mathbf{P}(B) \\ &= \mathbf{P}(A \mid C)\mathbf{P}(B \mid C),\end{aligned}$$

so A and B are conditionally independent given C. In the preceding calculation, the first equality uses the definition of conditional probabilities; the second uses the assumed independence; the fourth uses the independence of A from C, and of B from C.

Problem 45.* Assume that the events A_1, A_2, A_3, A_4 are independent and that $\mathbf{P}(A_3 \cap A_4) > 0$. Show that

$$\mathbf{P}(A_1 \cup A_2 \mid A_3 \cap A_4) = \mathbf{P}(A_1 \cup A_2).$$

Solution. We have

$$\mathbf{P}(A_1 \mid A_3 \cap A_4) = \frac{\mathbf{P}(A_1 \cap A_3 \cap A_4)}{\mathbf{P}(A_3 \cap A_4)} = \frac{\mathbf{P}(A_1)\mathbf{P}(A_3)\mathbf{P}(A_4)}{\mathbf{P}(A_3)\mathbf{P}(A_4)} = \mathbf{P}(A_1).$$

We similarly obtain $\mathbf{P}(A_2 \mid A_3 \cap A_4) = \mathbf{P}(A_2)$ and $\mathbf{P}(A_1 \cap A_2 \mid A_3 \cap A_4) = \mathbf{P}(A_1 \cap A_2)$, and finally,

$$\begin{aligned}\mathbf{P}(A_1 \cup A_2 \mid A_3 \cap A_4) &= \mathbf{P}(A_1 \mid A_3 \cap A_4) + \mathbf{P}(A_2 \mid A_3 \cap A_4) - \mathbf{P}(A_1 \cap A_2 \mid A_3 \cap A_4) \\ &= \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1 \cap A_2) \\ &= \mathbf{P}(A_1 \cup A_2).\end{aligned}$$

Problem 46.* **Laplace's rule of succession.** Consider $m + 1$ boxes with the kth box containing k red balls and $m - k$ white balls, where k ranges from 0 to m. We choose a box at random (all boxes are equally likely) and then choose a ball at random from that box, n successive times (the ball drawn is replaced each time, and a new ball is selected independently). Suppose a red ball was drawn each of the n times. What is the probability that if we draw a ball one more time it will be red? Estimate this probability for large m.

Solution. We want to find the conditional probability $\mathbf{P}(E \mid R_n)$, where E is the event of a red ball drawn at time $n + 1$, and R_n is the event of a red ball drawn each of the n preceding times. Intuitively, the consistent draw of a red ball indicates that a box with a high percentage of red balls was chosen, so we expect that $\mathbf{P}(E \mid R_n)$ is closer to 1 than to 0. In fact, Laplace used this example to calculate the probability that the sun will rise tomorrow given that it has risen for the preceding 5,000 years. (It is not clear how serious Laplace was about this calculation, but the story is part of the folklore of probability theory.)

We have

$$\mathbf{P}(E \mid R_n) = \frac{\mathbf{P}(E \cap R_n)}{\mathbf{P}(R_n)},$$

and by using the total probability theorem, we obtain

$$\mathbf{P}(R_n) = \sum_{k=0}^{m} \mathbf{P}(k\text{th box chosen}) \left(\frac{k}{m}\right)^n = \frac{1}{m+1} \sum_{k=0}^{m} \left(\frac{k}{m}\right)^n,$$

$$\mathbf{P}(E \cap R_n) = \mathbf{P}(R_{n+1}) = \frac{1}{m+1} \sum_{k=0}^{m} \left(\frac{k}{m}\right)^{n+1}.$$

For large m, we can view $\mathbf{P}(R_n)$ as a piecewise constant approximation to an integral:

$$\mathbf{P}(R_n) = \frac{1}{m+1} \sum_{k=0}^{m} \left(\frac{k}{m}\right)^n \approx \frac{1}{(m+1)m^n} \int_0^m x^n \, dx = \frac{1}{(m+1)m^n} \cdot \frac{m^{n+1}}{n+1} \approx \frac{1}{n+1}.$$

Similarly,

$$\mathbf{P}(E \cap R_n) = \mathbf{P}(R_{n+1}) \approx \frac{1}{n+2},$$

so that

$$\mathbf{P}(E \mid R_n) \approx \frac{n+1}{n+2}.$$

Thus, for large m, drawing a red ball one more time is almost certain when n is large.

Problem 47.* Binomial coefficient formula and the Pascal triangle.

(a) Use the definition of $\binom{n}{k}$ as the number of distinct n-toss sequences with k heads, to derive the recursion suggested by the so called Pascal triangle, given in Fig. 1.20.

(b) Use the recursion derived in part (a) and induction, to establish the formula

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

Solution. (a) Note that n-toss sequences that contain k heads (for $0 < k < n$) can be obtained in two ways:

(1) By starting with an $(n-1)$-toss sequence that contains k heads and adding a tail at the end. There are $\binom{n-1}{k}$ different sequences of this type.

(2) By starting with an $(n-1)$-toss sequence that contains $k-1$ heads and adding a head at the end. There are $\binom{n-1}{k-1}$ different sequences of this type.

$$
\begin{array}{ccccccccc}
 & & & & \binom{0}{0} & & & & \\
 & & & \binom{1}{0} & & \binom{1}{1} & & & \\
 & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\
 & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \\
 & & & & \cdots\cdots\cdots & & & &
\end{array}
\qquad
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
 & & & & \cdots\cdots\cdots & & & &
\end{array}
$$

Figure 1.20: Sequential calculation method of the binomial coefficients using the Pascal triangle. Each term $\binom{n}{k}$ in the triangular array on the left is computed and placed in the triangular array on the right by adding its two neighbors in the row above it (except for the boundary terms with $k = 0$ or $k = n$, which are equal to 1).

Thus,

$$\binom{n}{k} = \begin{cases} \binom{n-1}{k-1} + \binom{n-1}{k}, & \text{if } k = 1, 2, \ldots, n-1, \\ 1, & \text{if } k = 0, n. \end{cases}$$

This is the formula corresponding to the Pascal triangle calculation, given in Fig. 1.20.

(b) We now use the recursion from part (a), to demonstrate the formula

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!},$$

by induction on n. Indeed, we have from the definition $\binom{1}{0} = \binom{1}{1} = 1$, so for $n = 1$ the above formula is seen to hold as long as we use the convention $0! = 1$. If the formula holds for each index up to $n - 1$, we have for $k = 1, 2, \ldots, n-1$,

$$\begin{aligned}
\binom{n}{k} &= \binom{n-1}{k-1} + \binom{n-1}{k} \\
&= \frac{(n-1)!}{(k-1)!\,(n-1-k+1)!} + \frac{(n-1)!}{k!\,(n-1-k)!} \\
&= \frac{k}{n} \cdot \frac{n!}{k!\,(n-k)!} + \frac{n-k}{n} \cdot \frac{n!}{k!\,(n-k)!} \\
&= \frac{n!}{k!\,(n-k)!},
\end{aligned}$$

and the induction is complete.

Problem 48.* The Borel-Cantelli lemma. Consider an infinite sequence of trials. The probability of success at the ith trial is some positive number p_i. Let N be the event that there is no success, and let I be the event that there is an infinite number of successes.

(a) Assume that the trials are independent and that $\sum_{i=1}^{\infty} p_i = \infty$. Show that $\mathbf{P}(N) = 0$ and $\mathbf{P}(I) = 1$.

(b) Assume that $\sum_{i=1}^{\infty} p_i < \infty$. Show that $\mathbf{P}(I) = 0$.

Solution. (a) The event N is a subset of the event that there were no successes in the first n trials, so that

$$\mathbf{P}(N) \leq \prod_{i=1}^{n}(1 - p_i).$$

Taking logarithms,

$$\log \mathbf{P}(N) \leq \sum_{i=1}^{n} \log(1 - p_i) \leq \sum_{i=1}^{n}(-p_i).$$

Taking the limit as n tends to infinity, we obtain $\log \mathbf{P}(N) = -\infty$, or $\mathbf{P}(N) = 0$.

Let now L_n be the event that there is a finite number of successes and that the last success occurs at the nth trial. We use the already established result $\mathbf{P}(N) = 0$, and apply it to the sequence of trials after trial n, to obtain $\mathbf{P}(L_n) = 0$. The event I^c (finite number of successes) is the union of the disjoint events L_n, $n \geq 1$, and N, so that

$$\mathbf{P}(I^c) = \mathbf{P}(N) + \sum_{n=1}^{\infty} \mathbf{P}(L_n) = 0,$$

and $\mathbf{P}(I) = 1$.

(b) Let S_i be the event that the ith trial is a success. Fix some number n and for every $i > n$, let F_i be the event that the first success after time n occurs at time i. Note that $F_i \subset S_i$. Finally, let A_n be the event that there is at least one success after time n. Note that $I \subset A_n$, because an infinite number of successes implies that there are successes subsequent to time n. Furthermore, the event A_n is the union of the disjoint events F_i, $i > n$. Therefore,

$$\mathbf{P}(I) \leq \mathbf{P}(A_n) = \mathbf{P}\left(\bigcup_{i=n+1}^{\infty} F_i\right) = \sum_{i=n+1}^{\infty} \mathbf{P}(F_i) \leq \sum_{i=n+1}^{\infty} \mathbf{P}(S_i) = \sum_{i=n+1}^{\infty} p_i.$$

We take the limit of both sides as $n \to \infty$. Because of the assumption $\sum_{i=1}^{\infty} p_i < \infty$, the right-hand side converges to zero. This implies that $\mathbf{P}(I) = 0$.

SECTION 1.6. Counting

Problem 49. De Méré's puzzle. A six-sided die is rolled three times independently. Which is more likely: a sum of 11 or a sum of 12? (This question was posed by the French nobleman de Méré to his friend Pascal in the 17th century.)

Problem 50. The birthday problem. Consider n people who are attending a party. We assume that every person has an equal probability of being born on any day during the year. independent of everyone else, and ignore the additional complication presented by leap years (i.e., assume that nobody is born on February 29). What is the probability that each person has a distinct birthday?

Problem 51. An urn contains m red and n white balls.

(a) We draw two balls randomly and simultaneously. Describe the sample space and calculate the probability that the selected balls are of different color, by using two approaches: a counting approach based on the discrete uniform law, and a sequential approach based on the multiplication rule.

(b) We roll a fair 3-sided die whose faces are labeled 1,2,3, and if k comes up. we remove k balls from the urn at random and put them aside. Describe the sample space and calculate the probability that all of the balls drawn are red. using a divide-and-conquer approach and the total probability theorem.

Problem 52. We deal from a well-shuffled 52-card deck. Calculate the probability that the 13th card is the first king to be dealt.

Problem 53. Ninety students, including Joe and Jane, are to be split into three classes of equal size, and this is to be done at random. What is the probability that Joe and Jane end up in the same class?

Problem 54. Twenty distinct cars park in the same parking lot every day. Ten of these cars are US-made. while the other ten are foreign-made. The parking lot has exactly twenty spaces. all in a row. so the cars park side by side. However. the drivers have varying schedules. so the position any car might take on a certain day is random.

(a) In how many different ways can the cars line up?

(b) What is the probability that on a given day, the cars will park in such a way that they alternate (no two US-made are adjacent and no two foreign-made are adjacent)?

Problem 55. Eight rooks are placed in distinct squares of an 8×8 chessboard, with all possible placements being equally likely. Find the probability that all the rooks are safe from one another, i.e.. that there is no row or column with more than one rook.

Problem 56. An academic department offers 8 lower level courses: $\{L_1, L_2, \ldots, L_8\}$ and 10 higher level courses: $\{H_1, H_2, \ldots, H_{10}\}$. A valid curriculum consists of 4 lower level courses. and 3 higher level courses.

(a) How many different curricula are possible?

(b) Suppose that $\{H_1, \ldots, H_5\}$ have L_1 as a prerequisite, and $\{H_6, \ldots H_{10}\}$ have L_2 and L_3 as prerequisites. i.e.. any curricula which involve, say, one of $\{H_1, \ldots, H_5\}$ must also include L_1. How many different curricula are there?

Problem 57. How many 6-word sentences can be made using each of the 26 letters of the alphabet exactly once? A word is defined as a nonempty (possibly jibberish) sequence of letters.

Problem 58. We draw the top 7 cards from a well-shuffled standard 52-card deck. Find the probability that:

(a) The 7 cards include exactly 3 aces.

(b) The 7 cards include exactly 2 kings.

(c) The probability that the 7 cards include exactly 3 aces. or exactly 2 kings, or both.

Problem 59. A parking lot contains 100 cars, k of which happen to be lemons. We select m of these cars at random and take them for a test drive. Find the probability that n of the cars tested turn out to be lemons.

Problem 60. A well-shuffled 52-card deck is dealt to 4 players. Find the probability that each of the players gets an ace.

Problem 61.* **Hypergeometric probabilities.** An urn contains n balls, out of which m are red. We select k of the balls at random. without replacement (i.e., selected balls are not put back into the urn before the next selection). What is the probability that i of the selected balls are red?

Solution. The sample space consists of the $\binom{n}{k}$ different ways that we can select k out of the available balls. For the event of interest to occur. we have to select i out of the m red balls, which can be done in $\binom{m}{i}$ ways, and also select $k-i$ out of the $n-m$ balls that are not red, which can be done in $\binom{n-m}{k-i}$ ways. Therefore, the desired probability is

$$\frac{\binom{m}{i}\binom{n-m}{k-i}}{\binom{n}{k}},$$

for $i \geq 0$ satisfying $i \leq m$, $i \leq k$, and $k-i \leq n-m$. For all other i, the probability is zero.

Problem 62.* **Correcting the number of permutations for indistinguishable objects.** When permuting n objects, some of which are indistinguishable, different permutations may lead to indistinguishable object sequences, so the number of distinguishable object sequences is less than $n!$. For example, there are six permutations of the letters A, B, and C:

ABC. ACB, BAC, BCA, CAB, CBA,

but only three distinguishable sequences that can be formed using the letters A, D, and D:

ADD, DAD. DDA.

(a) Suppose that k out of the n objects are indistinguishable. Show that the number of distinguishable object sequences is $n!/k!$.

(b) Suppose that we have r types of indistinguishable objects, and for each i, k_i objects of type i. Show that the number of distinguishable object sequences is

$$\frac{n!}{k_1!\,k_2!\cdots k_r!}.$$

Solution. (a) Each one of the $n!$ permutations corresponds to $k!$ duplicates which are obtained by permuting the k indistinguishable objects. Thus, the $n!$ permutations can be grouped into $n!/k!$ groups of $k!$ indistinguishable permutations that result in the same object sequence. Therefore, the number of distinguishable object sequences is $n!/k!$. For example, the three letters A, D, and D give the $3! = 6$ permutations

$$\text{ADD, ADD, DAD, DDA, DAD, DDA,}$$

obtained by replacing B and C by D in the permutations of A, B, and C given earlier. However, these 6 permutations can be divided into the $n!/k! = 3!/2! = 3$ groups

$$\{\text{ADD}, \text{ADD}\}, \ \{\text{DAD}, \text{DAD}\}, \ \{\text{DDA}, \text{DDA}\},$$

each having $k! = 2! = 2$ indistinguishable permutations.

(b) One solution is to extend the argument in (a) above: for each object type i, there are $k_i!$ indistinguishable permutations of the k_i objects. Hence, each permutation belongs to a group of $k_1!\,k_2!\cdots k_r!$ indistinguishable permutations, all of which yield the same object sequence.

An alternative argument goes as follows. Choosing a distinguishable object sequence is the same as starting with n slots and for each i, choosing the k_i slots to be occupied by objects of type i. This is the same as partitioning the set $\{1, \ldots, n\}$ into groups of size $k_1, \ldots, k_r$, and the number of such partitions is given by the multinomial coefficient.