INTRODUCTION TO PROBABILITY

by

Dimitri P. Bertsekas and John N. Tsitsiklis

CHAPTER 1: ADDITIONAL PROBLEMS

Last updated: September 12, 2005

SECTION 1.1. Sets.

Problem 1. We are given that $P(A) = 0.55$, $P(B^c) = 0.35$, and $P(A \cup B) = 0.75$. Determine $P(B)$ and $P(A \cap B)$.

Problem 2. Let A and B be two sets. Under what conditions is the set $A \cap (A \cup B)^c$ empty?

Problem 3. Let A and B be two sets.

(a) Show that $(A^c \cap B^c)^c = A \cup B$ and $(A^c \cup B^c)^c = A \cap B$.

(b) Consider rolling a six-sided die once. Let A be the set of outcomes where an odd number comes up. Let B be the set of outcomes where a 1 or a 2 comes up. Calculate the sets on both sides of the equalities in part (a), and verify that the equalities hold.

Problem 4. Let A and B be two sets with a finite number of elements. Show that the number of elements in $A \cap B$ plus the number of elements in $A \cup B$ is equal to the number of elements in A plus the number of elements in B.

SECTION 1.2. Probabilistic Models

Problem 5. We are given that $\mathbf{P}(A^c) = 0.6$, $\mathbf{P}(B) = 0.3$, and $\mathbf{P}(A \cap B) = 0.2$. Determine $\mathbf{P}(A \cup B)$.

Problem 6. We roll a four-sided die once and then we roll it as many times as is necessary to obtain a different face than the one obtained in the first roll. Let the outcome of the experiment be (r_1, r_2) where r_1 and r_2 are the results of the first and the last rolls, respectively. Assume that all possible outcomes have equal probability. Find the probability that:

(a) r_1 is even.

(b) Both r_1 and r_2 are even.

(c) $r_1 + r_2 < 5$.

Problem 7. A magical four-sided die is rolled twice. Let S be the sum of the results of the two rolls. We are told that the probability that $S = k$ is proportional to k, for $k = 2, 3, \ldots, 8$, and that all possible ways that a given sum k can arise are equally likely. Construct an appropriate probabilistic model and find the probability of getting doubles.

Problem 8. You enter a special kind of chess tournament, whereby you play one game with each of three opponents, but you get to choose the order in which you play your opponents. You win the tournament if you win two games in a row. You know your probability of a win against each of the three opponents. What is your probability of winning the tournament, assuming that you choose the optimal order of playing the opponents?

Problem 9. Alice and Bob each choose at random a number between zero and two. We assume a uniform probability law under which the probability of an event is proportional to its area. Consider the following events:

A: The magnitude of the difference of the two numbers is greater than 1/3.

B: At least one of the numbers is greater than 1/3.

C: The two numbers are equal.

D: Alice's number is greater than 1/3.

Find the probabilities $\mathbf{P}(A)$, $\mathbf{P}(B)$, $\mathbf{P}(A \cap B)$, $\mathbf{P}(C)$, $\mathbf{P}(D)$, $\mathbf{P}(A \cap D)$.

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. Show the following generalizations of the formula

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

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

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

(b) Let $A_1, A_2, \ldots, A_n$ be events. Then

$$\mathbf{P}\left(\cup_{k=1}^{n} A_k\right) = \mathbf{P}(A_1) + \mathbf{P}(A_1^c \cap A_2) + \mathbf{P}(A_1^c \cap A_2^c \cap A_3) + \cdots + \mathbf{P}(A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n).$$

SECTION 1.3. Conditional Probability

Problem 12. The disc containing the only copy of your thesis just got corrupted, and the disk got mixed up with three other corrupted discs that were lying around. It is equally likely that any of the four discs holds the corrupted remains of your thesis.

Your computer expert friend offers to have a look, and you know from past experience that his probability of finding your thesis from any disc is 0.4 (assuming the thesis is there). Given that he searches on disc 1 but cannot find your thesis, what is the probability that your thesis is on disc i, for $i = 1, 2, 3, 4$?

Problem 13. A person has forgotten the last digit of a telephone number, so he dials the number with the last digit randomly chosen. How many times does he have to dial (not counting repetitions) in order that the probability of dialing the correct number is more than 0.5.

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 were rolled.

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

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

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

SECTION 1.4. Total Probability Theorem and Bayes' Rule

Problem 15. A new test has been developed to determine whether a given student is overstressed. This test is 95% accurate if the student is not overstressed, but only 85% accurate if the student is in fact overstressed. It is known that 99.5% of all students are overstressed. Given that a particular student tests negative for stress, what is the probability that the test results are correct, and that this student is not overstressed?

Problem 16. A hiker starts by taking one of n available trails, denoted $1, 2, \ldots, n$. An hour into the hike, trail i subdivides into $1 + i$ subtrails, only one of which leads to the hiker's destination. The hiker has no map and makes random choices of trail and subtrail. What is the probability of reaching the destination?

Problem 17. Alice and Bob have $2n + 1$ coins, each with probability of a head equal to $1/2$. Bob tosses $n + 1$ coins, while Alice tosses the remaining n coins. Show that the probability that after all the coins have been tossed, Bob will have gotten more heads than Alice is $1/2$.

Problem 18. A magnetic tape storing information in binary form has been corrupted, so it can only be read with bit errors. The probability that you correctly detect a 0 is 0.9, while the probability that you correctly detect a 1 is 0.85. Each digit is a 1 or a 0 with equal probability. Given that you read a 1, what is the probability that this is a correct reading?

SECTION 1.5. Independence

Problem 19. An internet access provider (IAP) owns two servers. Each server has a 50% chance of being "down" independently of the other. Fortunately, only one server is necessary to allow the IAP to provide service to its customers, i.e., only one server is needed to keep the IAP's system up. Suppose a customer tries to access the internet on four different occasions, which are sufficiently spaced apart in time, so that we may assume that the states of the system corresponding to these four occasions are independent. What is the probability that the customer will only be able to access the internet on 3 out of the 4 occasions?

Problem 20. A peculiar six-sided die has uneven faces. In particular, the faces showing 1 or 6 are 1×1.5 inches, the faces showing 2 or 5 are 1×0.4 inches, and the faces showing 3 or 4 are 0.4×1.5 inches. Assume that the probability of a particular face coming up is proportional to its area. We independently roll the die twice. What is the probability that we get doubles?

Problem 21. A parking lot consists of a single row containing n parking spaces ($n \geq 2$). Mary arrives when all spaces are free. Tom is the next person to arrive. Each person makes an equally likely choice among all available spaces at the time of arrival. Describe the sample space. Obtain $\mathbf{P}(A)$, the probability the parking spaces selected by Mary and Tom are at most 2 spaces apart.

Problem 22. We are given three coins. The first coin is a fair coin painted blue on the head side and white on the tail side. The other two coins are biased so that the probability of a head is p. They are painted blue on the tail side and red on the head side. Two of the three coins are to be selected at random and tossed. Describe the outcomes in the sample space. It was experimentally determined that the probability that the sides that land face up are of the same color is $29/96$. What are the possible values of p?

Problem 23. A company is interviewing potential employees. Suppose that each candidate is either qualified, or unqualified with given probabilities q and $1-q$, respectively. The company tries to determine a candidate's qualifications by asking 20 true-false questions. A qualified candidate has probability p of answering a question correctly, while an unqualified candidate has a probability p of answering incorrectly. The answers to different questions are assumed to be independent. If the company considers anyone with at least 15 correct answers qualified, and everyone else unqualified, give a formula for the probability that the 20 questions will correctly identify someone to be qualified or unqualified.

Problem 24. A particular jury consists of 7 jurors. Each juror has a 0.2 chance of making the wrong decision, independently of the others. If the jury reaches a decision by majority rule, what is the probability that it will reach a wrong decision?

Problem 25. Three persons roll a fair n-sided die once. Let A_{ij} be the event that person i and person j roll the same face. Show that the events A_{12}, A_{13}, and A_{23} are pairwise independent but are not independent.

Problem 26. Calculating the odds. Let A be an event such that $0 < \mathbf{P}(A) < 1$. The *odds in favor* of A are defined to be

$$\mathbf{O}(A) = \frac{\mathbf{P}(A)}{\mathbf{P}(A^c)},$$

while the *odds against* A are defined to be the reciprocal of $\mathbf{O}(A)$. [To connect the term "odds" with its common usage, note for example that if the probability that a given horse wins a race at the track is $1/3$, the odds against the horse winning are 2 to 1. A "fair" racetrack would then pay \$2 for every \$1 bet on the horse (plus the original \$1 bet), if the horse wins; "fair" here means that the racetrack would break even on the average – this will become more precise in Chapter 2, when we will discuss the notion of expected value.] This problem deals with a formula for calculating "conditional odds," that is, odds based on some partial information. If A and B are events with $\mathbf{P}(A) > 0$ and $\mathbf{P}(B) > 0$, the odds in favor of A given B are defined as

$$\mathbf{O}(A \mid B) = \frac{\mathbf{P}(A \mid B)}{\mathbf{P}(A^c \mid B)}.$$

Show that

$$\mathbf{O}(A \mid B) = \mathbf{L}(B \mid A)\mathbf{O}(A),$$

where $\mathbf{L}(B \mid A)$ is the so called *likelihood ratio* of B given A, defined as

$$\mathbf{L}(B \mid A) = \frac{\mathbf{P}(B \mid A)}{\mathbf{P}(B \mid A^c)}.$$

Problem 27. Hypothesis testing. May B. Lucky is a compulsive gambler who is convinced that on any given day she is either "lucky," in which case she wins each red/black bet she makes in the roulette with probability $p_L > 1/2$, or she is "unlucky," in which case she wins each red/black bet she makes in the roulette with probability $p_U < 1/2$. May visits the casino every day, and believes that she knows the *a priori* probability that any one given visit is a "lucky" one (i.e., corresponds to p_L rather than p_U). To improve her chances, May adopts a system whereby she estimates on-line whether she is lucky or unlucky on a given day, by keeping a running count of the numbers of bets that she wins and loses. In particular, she continues to play until the conditional odds in favor of the event {lucky on the current day}, given the number of wins and losses so far, fall below a certain threshold (see the preceding problem). As soon as this happens, she stops playing. Provide a simple algorithm for updating May's conditional odds with each play. *Note*: This example is typical of reasoning in *sequential hypothesis testing systems*, where the probability of correctness of a certain hypothesis, given some evidence, is calculated and sequentially updated.

Problem 28. Let A and B be events such that $A \subset B$. Can A and B be independent?

Problem 29. We are told that events A and B are independent. In addition, events A and C are independent. Is it true that A is independent of $B \cup C$? Provide a proof or counterexample to support your answer.

Problem 30. Suppose that A, B, and C are independent. Use the definition of independence to show that A and $B \cup C$ are independent.

SECTION 1.6. Counting

Problem 31. A parking lot contains 100 cars that all look quite nice from the outside. However, K of these cars happen to be lemons. The number K is known to lie in the range $\{0, 1, \ldots, 9\}$, with all values equally likely.

(a) We testdrive 20 distinct cars chosen at random, and to our pleasant surprise, none of them turns out to be a lemon. Given this knowledge, what is the probability that $K = 0$?

(b) Repeat part (a) when the 20 cars are chosen with replacement; that is, at each testdrive, each car is equally likely to be selected, including those that were selected earlier.

Problem 32. A candy factory has an endless supply of red, orange, yellow, green, blue, and violet jelly beans. The factory packages the jelly beans into jars of 100 jelly beans each. One possible color distribution, for example, is a jar of 58 red, 22 yellow, and 20 green jelly beans. As a marketing gimmick, the factory guarantees that no two jars have the same color distribution. What is the maximum number of jars the factory can produce?

Problem 33. A fair six-sided die is rolled repeatedly and independently. Let A_n be the event of rolling n sixes in $6n$ rolls, and let B_n be the event of rolling n or more sixes in $6n$ rolls.

(a) Does $\mathbf{P}(A_n)$ and $\mathbf{P}(B_n)$ change with n?

(b) Use a computer to investigate what happens to $\mathbf{P}(A_n)$ and $\mathbf{P}(B_n)$ as n becomes very large.

Problem 34. Consider three independent rolls of a fair six-sided die.

(a) What is the probability that the sum of the three rolls is 11?

(b) What is the probability that the sum of the three rolls is 12?

(c) In the seventeenth century, Galileo explained the experimental observation that a sum of 10 is more frequent than a sum of 9, even though both 10 and 9 can be obtained in six distinct ways. Can you retrace Galileo's thinking?

Problem 35. Consider a backgammon match with 25 games, each of which can have one of two outcomes: win (1 point), or loss (0 points). Find the number of all possible distinct score sequences under the following alternative assumptions.

(a) All 25 games are played.

(b) The match is stopped when one player reaches 13 points.

Problem 36. Count the number of distinguishable ways in which you can arrange the letters in the words:

(a) children

(b) bookkeeper