Probability theory

Introduction

We often hear statements of the following kind: "It is likely to rain today", "I have a fair chance of passing this course", "There is an even chance that a coin will come up heads", etc. In each case our statement refers to a situation in which we are not certain of the outcome, but we express some degree of confidence that our prediction will be verified. The theory of probability provides a mathematical framework for such assertions.

Consider an experiment whose outcome is not known. Suppose that someone makes an assertion p about the outcome of the experiment, and we want to assign a probability to p. When statement p is considered in isolation, we usually find no natural assignment of probabilities. Rather, we look for a method of assigning probabilities to all conceivable statements concerning the outcome of the experiment. At first this might seem to be a hopeless task, since there is no end to the statements we can make about the experiment. However we are aided by a basic principle:

Fundamental assumption. Any two equivalent statements will be assigned the same probability. As long as there are a finite number of logical possibilities, there are only a finite number of truth sets, and hence the process of assigning probabilities is a finite one. We proceed in three steps: (1) we first determine $\mathcal{U}$, the possibility set, that is, the set of all logical possibilities, (2) to each subset X of $\mathcal{U}$ we assign a number called the measure $m(X)$, (3) to each statement p we assign $m(P)$, the measure of its truth set, as a probability. The probability of statement p is denoted by $\Pr[p]$.

The first step, that of determining the set of logical possibilities, is one that we considered in the previous chapters. It is important to recall that there is no unique method for analyzing logical possibilities. In a given problem we may arrive at a very fine or a very rough analysis of possibilities, causing $\mathcal{U}$ to have many or few elements.

Having chosen $\mathcal{U}$, the next step is to assign a number to each subset X of $\mathcal{U}$, which will in turn be taken to be the probability of any statement having truth set X. We do this in the following way.

Assignment of a measure. Assign a positive number (weight) to each element of $\mathcal{U}$, so that the sum of the weights assigned is 1. Then the measure of a set is the sum of the weights of its elements. The measure of the set $\emptyset$ is 0.

In applications of probability to scientific problems, the analysis of the logical possibilities and the assignment of measures may depend upon factual information and hence can best be done by the scientist making the application.

Once the weights are assigned, to find the probability of a particular statement we must find its truth set and find the sum of the weights assigned to elements of the truth set. This problem, which might seem easy, can often involve considerable mathematical difficulty. The development of techniques to solve this kind of problem is the main task of probability theory.

Example 4.1 An ordinary die is thrown. What is the probability that the number which turns up is less than four? Here the possibility set is $\mathcal{U} = \{1, 2, 3, 4, 5, 6\}$. The symmetry of the die suggests that each face should have the same probability of turning up. To make this so, we assign weight $\frac{1}{6}$ to each of the outcomes. The truth set of the statement "The number which turns up is less than four" is $\{1, 2, 3\}$. Hence the probability of this statement is $\frac{3}{6} = \frac{1}{2}$, the sum of the weights of the elements in its truth set. ◇

Example 4.2 A gambler attends a race involving three horses A, B, and C. He or she feels that A and B have the same chance of winning but that A (and hence also B) is twice as likely to win as C is. What is the probability that A or C wins? We take as $\mathcal{U}$ the set $\{\mathrm{A}, \mathrm{B}, \mathrm{C}\}$. If we were to assign weight a to the outcome C, then we would assign weight $2a$ to each of the outcomes A and B. Since the sum of the weights must be 1, we have $2a + 2a + a = 1$, or $a = \frac{1}{5}$. Hence we assign weights $\frac{2}{5}$, $\frac{2}{5}$, $\frac{1}{5}$ to the outcomes A, B, and C, respectively. The truth set of the statement "Horse A or C wins" is $\{A, C\}$. The sum of the weights of the elements of this set is $\frac{2}{5} + \frac{1}{5} = \frac{3}{5}$. Hence the probability that A or C wins is $\frac{3}{5}$. $\diamond$

Exercises

1. Assume that there are n possibilities for the outcome of a given experiment. How should the weights be assigned if it is desired that all outcomes be assigned the same weight?

2. Let $\mathcal{U} = \{a, b, c\}$. Assign weights to the three elements so that no two have the same weight, and find the measures of the eight subsets of $\mathcal{U}$.

3. In an election Jones has probability $\frac{1}{2}$ of winning, Smith has probability $\frac{1}{3}$, and Black has probability $\frac{1}{6}$.

   (a) Construct $\mathcal{U}$.

   (b) Assign weights.

   (c) Find the measures of the eight subsets.

   (d) Give a pair of nonequivalent predictions which have the same probability.

4. Give the possibility set $\mathcal{U}$ for each of the following experiments.

   (a) An election between candidates A and B is to take place.

   (b) A number from 1 to 5 is chosen at random.

   (c) A two-headed coin is thrown.

   (d) A student is asked for the day of the year on which his or her birthday falls.

5. For which of the cases in Exercise 4 might it be appropriate to assign the same weight to each outcome?

6. Suppose that the following probabilities have been assigned to the possible results of putting a penny in a certain defective peanut-vending machine: The probability that nothing comes out is $\frac{1}{2}$. The probability that either you get your money back or you get peanuts (but not both) is $\frac{1}{3}$.

(a) What is the probability that you get your money back and also get peanuts?

[Ans. $\frac{1}{6}$.]

(b) From the information given, is it possible to find the probability that you get peanuts?

[Ans. No.]

7. A die is loaded in such a way that the probability of each face is proportional to the number of dots on that face. (For instance, a 6 is three times as probable as a 2.) What is the probability of getting an even number in one throw?

[Ans. $\frac{4}{7}$.]

8. If a coin is thrown three times, list the eight possibilities for the outcomes of the three successive throws. A typical outcome can be written (HTH). Determine a probability measure by assigning an equal weight to each outcome. Find the probabilities of the following statements.

(a) r: The number of heads that occur is greater than the number of tails.

[Ans. $\frac{1}{2}$.]

(b) s: Exactly two heads occur.

[Ans. $\frac{5}{8}$.]

(c) t: The same side turns up on every throw.

[Ans. $\frac{1}{4}$.]

9. For the statements given in Exercise 8, which of the following equalities are true?

(a) $\Pr[r \vee s] = \Pr[r] + \Pr[s]$

(b) $\Pr[s \vee t] = \Pr[s] + \Pr[t]$

(c) $\Pr[r \vee \neg r] = \Pr[r] + \Pr[\neg r]$

(d) $\Pr[r \vee t] = \Pr[r] + \Pr[t]$

10. Which of the following pairs of statements (see Exercise 8) are inconsistent? (Recall that two statements are inconsistent if their truth sets have no element in common.)

    (a) r, s

    [Ans. consistent.]

    (b) s, t

    [Ans. inconsistent.]

    (c) $r, \neg r$

    [Ans. inconsistent.]

    (d) r, t

    [Ans. consistent.]

11. State a theorem suggested by Exercises 9 and 10.

12. An experiment has three possible outcomes, a, b, and c. Let p be the statement "the outcome is a or b", and q be the statement "the outcome is b or c". Assume that weights have been assigned to the three outcomes so that $\Pr[p] = \frac{2}{3}$ and $\Pr[q] = \frac{5}{6}$. Find the weights.

    [Ans. $\frac{1}{6}, \frac{1}{2}, \frac{1}{3}$.]

13. Repeat Exercise 12 if $\Pr[p] = \frac{1}{2}$ and $\Pr[q] = \frac{3}{8}$.

Properties of a probability measure

Before studying special probability measures, we shall consider some general properties of such measures which are useful in computations and in the general understanding of probability theory.

Three basic properties of a probability measure are

(A) $m(X) = 0$ if and only if $X = \emptyset$.

(B) $0 < m(X) < 1$ for any set X.

(C) For two sets X and Y,

$$m(X \cup Y) = m(X) + m(Y)$$

if and only if X and Y are disjoint, i.e., have no elements in common.

The proofs of properties (A) and (B) are left as an exercise (see Exercise 16). We shall prove (C). We observe first that $m(X) + m(Y)$ is the sum of the weights of the elements of X added to the sum of the weights of Y. If X and Y are disjoint, then the weight of every element of $X \cup Y$ is added once and only once, and hence $m(X) + m(Y) = m(X \cup Y)$. Assume now that X and Y are not disjoint. Here the weight of every element contained in both X and Y, i.e., in $X \cap Y$, is added twice in the sum $m(X) + m(Y)$. Thus this sum is greater than $m(X \cup Y)$ by an amount $m(X \cap Y)$. By (A) and (B), if $X \cap Y$ is not the empty set, then $m(X \cap Y) > 0$. Hence in this case we have $m(X) + m(Y) > m(X \cap Y)$. Thus if X and Y are not disjoint, the equality in (C) does not hold. Our proof shows that in general we have

(C′) For any two sets X and Y, $m(X \cup Y) = m(X) + m(Y) - m(X \cap Y)$.

Since the probabilities for statements are obtained directly from the probability measure $m(X)$, any property of $m(X)$ can be translated into a property about the probability of statements. For example, the above properties become, when expressed in terms of statements,

(a) $\Pr[p] = 0$ if and only if p is logically false.

(b) $0 < \Pr[p] < 1$ for any statement p.

(c) The equality

$$\Pr[p \vee q] = \Pr[p] + \Pr[q]$$

holds and only if p and q are inconsistent.

(c′) For any two statements p and q,

$$\Pr[p \vee q] = \Pr[p] + \Pr[q] - \Pr[p \wedge q].$$

Another property of a probability measure which is often useful in computation is

(D) $m(\tilde{X}) = 1 - m(X)$,

or, in the language of statements,

(d) $\Pr[\neg p] = 1 - \Pr[p]$.

The proofs of (D) and (d) are left as an exercise (see Exercise 17).

It is important to observe that our probability measure assigns probability 0 only to statements which are logically false, i.e., which are false for every logical possibility. Hence, a prediction that such a statement will be true is certain to be wrong. Similarly, a statement is assigned probability 1 only if it is true in every case, i.e., logically true. Thus the prediction that a statement of this type will be true is certain to be correct. (While these properties of a probability measure seem quite natural, it is necessary, when dealing with infinite possibility sets, to weaken them slightly. We consider in this book only the finite possibility sets.)

We shall now discuss the interpretation of probabilities that are not 0 or 1. We shall give only some intuitive ideas that are commonly held concerning probabilities. While these ideas can be made mathematically more precise, we offer them here only as a guide to intuitive thinking.

Suppose that, relative to a given experiment, a statement has been assigned probability p. From this it is often inferred that if a sequence of such experiments is performed under identical conditions, the fraction of experiments which yield outcomes making the statement true would be approximately p. The mathematical version of this is the "law of large numbers" of probability theory (which will be treated in Section 4.10). In cases where there is no natural way to assign a probability measure, the probability of a statement is estimated experimentally. A sequence of experiments is performed and the fraction of the experiments which make the statement true is taken as the approximate probability for the statement.

A second and related interpretation of probabilities is concerned with betting. Suppose that a certain statement has been assigned probability p. We wish to offer a bet that the statement will in fact turn out to be true. We agree to give r dollars if the statement does not turn out to be true, provided that we receive s dollars if it does turn out to be true. What should r and s be to make the bet fair? If it were true that in a large number of such bets we would win s a fraction p of the times and lose r a fraction $1-p$ of the time, then our average winning per bet would be $sp-r(1-p)$. To make the bet fair we should make this average winning 0. This will be the case if $sp = r(1-p)$ or if $r/s = p/(1-p)$. Notice that this determines only the ratio of r and s. Such a ratio, written $r : s$, is said to give *odds* in favor of the statement.

Definition. The *odds* in favor of an outcome are $r : s$ (r *to* s), if the probability of the outcome is p, and $r/s = p/(1-p)$. Any two numbers having the required ratio may be used in place of r and s. Thus $6 : 4$ odds are the same as $3 : 2$ odds.

Example 4.3 Assume that a probability of $\frac{3}{4}$ has been assigned to a certain horse winning a race. Then the odds for a fair bet would be $\frac{3}{4} : \frac{1}{4}$. These odds could be equally well written as $3 : 1$, $6 : 2$ or $12 : 4$, etc. A fair bet would be to agree to pay \$3 if the horse loses and receive \$1 if the horse wins. Another fair bet would be to pay \$6 if the horse loses and win \$2 if the horse wins. ◇

Exercises

1. Let p and q be statements such that $\Pr[p \wedge q] = \frac{1}{4}$, $\Pr[\neg p] = \frac{1}{3}$, and $\Pr[q] = \frac{1}{2}$. What is $\Pr[p \vee q]$?

   [Ans. $\frac{11}{12}$.]

2. Using the result of Exercise 1, find $\Pr[\neg p \wedge \neg q]$.

3. Let p and q be statements such that $\Pr[p] = \frac{1}{2}$ and $\Pr[q] = \frac{2}{3}$. Are p and q consistent?

   [Ans. Yes.]

4. Show that, if $\Pr[p] + \Pr[q] > 1$, then p and q are consistent.

5. A student is worried about his or her grades in English and Art. The student estimates that the probability of passing English is .4, the probability of passing at least one course with probability .6, but that the probability of 1 of passing both courses is only .1? What is the probability that the student will pass Art?

   [Ans. .3.]

6. Given that a school has grades A, B, C, D, and F, and that a student has probability .9 of passing a course, and .6 of getting a grade lower than B, what is the probability that the student will get a C or D?

   [Ans. $\frac{1}{2}$.]

7. What odds should a person give on a bet that a six will turn up when a die is thrown?

8. Referring to Example 4.2, what odds should the man be willing to give for a bet that either A or B will come in first?

9. Prove that if the odds in favor of a given statement are $r : s$, then the probability that the statement will be true is $r/(r + s)$.

10. Using the result of Exercise 9 and the definition of "odds", show that if the odds are $r : s$ that a statement is true, then the odds are $s : r$ that it is false.

11. A gambler is willing to give 5 : 4 odds that the Dodgers will win the World Series. What must the probability of a Dodger victory be for this to be a fair bet?

[Ans. $\frac{5}{9}$.]

12. A statistician has found through long experience that if he or she washes the car, it rains the next day 85 per cent of the time. What odds should the statistician give that this will occur next time?

13. A gambler offers 1 : 3 odds that A will occur, 1 : 2 odds that B will occur. The gambler knows that A and B cannot both occur. What odds should he or she give that A or B will occur?

[Ans. 7 : 5.]

14. A gambler offers 3 : 1 odds that A will occur, 2 : 1 odds that B will occur. The gambler knows that A and B cannot both occur. What odds should he or she give that A or B will occur?

15. Show from the definition of a probability measure that $m(X) = 1$ if and only if $X = \mathcal{U}$.

16. Show from the definition of a probability measure that properties (A), (B) of the text are true.

17. Prove property (D) of the text. Why does property (d) follow from this property?

18. Prove that if R, S, and T are three sets that have no element in common,

$$m(R \cup S \cup T) = m(R) + m(S) + m(T).$$

19. If X and Y are two sets such that X is a subset of Y, prove that $m(X) \leq m(Y)$.

20. If p and q are two statements such that p implies q, prove that $\Pr[p] \leq \Pr[q]$.

21. Suppose that you are given n statements and each has been assigned a probability less than or equal to r. Prove that the probability of the disjunction of these statements is less than or equal to nr.

22. The following is an alternative proof of property (C′) of the text. Give a reason for each step.

    (a) $X \cup Y = (X \cap \tilde{Y}) \cup (X \cap Y) \cup (\tilde{X} \cap Y)$.

    (b) $m(X \cup Y) = m(X \cap \tilde{Y}) + m(X \cap Y) + m(\tilde{X} \cap Y)$.

    (c) $m(X \cup Y) = m(X) + m(Y) - m(X \cap Y)$.

23. If X, Y, and Z are any three sets, prove that, for any probability measure,

$$\begin{aligned} m(X \cup Y \cup Z) &= m(X) + m(Y) + m(Z) \\ &\quad - m(X \cap Y) - m(Y \cap Z) - m(X \cap Z) \\ &\quad + m(X \cap Y \cap Z). \end{aligned}$$

24. Translate the result of Exercise 23 into a result concerning three statements p, q, and r.

25. A man offers to bet "dollars to doughnuts" that a certain event will take place. Assuming that a doughnut costs a nickel, what must the probability of the event be for this to be a fair bet?

[Ans. $\frac{20}{21}$.]

26. Show that the inclusion-exclusion formula 3.2 is true is n is replaced by m. Apply this result to

$$\Pr[p_1 \vee p_2 \vee \ldots \vee p_n].$$

The equiprobable measure

We have already seen several examples where it was natural to assign the same weight to all possibilities in determining the appropriate probability measure. The probability measure determined in this manner is called the equiprobable measure. The measure of sets in the case of the equiprobable measure has a very simple form. In fact, if $\mathcal{U}$ has n elements and if the equiprobable measure has been assigned, then for any set X, $m(X)$ is r/n, where r is the number of elements in the set X. This is true since the weight of each element in X is $1/n$, and hence the sum of the weights of elements of X is r/n.

The particularly simple form of the equiprobable measure makes it easy to work with. In view of this, it is important to observe that a particular choice for the set of possibilities in a given situation may lead to the equiprobable measure, while some other choice will not. For example, consider the case of two throws of an ordinary coin. Suppose that we are interested in statements about the number of heads which occur. If we take for the possibility set the set $\mathcal{U} = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$ then it is reasonable to assign the same weight to each outcome, and we are led to the equiprobable measure. If, on the other hand, we were to take as possible outcomes the set $\mathcal{U} = \{\text{no H}, \text{one H}, \text{two H}\}$, it would not be natural to assign the same weight to each outcome, since one head can occur in two different ways, while each of the other possibilities can occur in only one way.

Example 4.4 Suppose that we throw two ordinary dice. Each die can turn up a number from 1 to 6; hence there are $6 \cdot 6$ possibilities. We assign weight $\frac{1}{36}$ to each possibility. A prediction that is true in j cases will then have probability $j/36$. For example, "The sum of the dice is 5" will be true if we get $1 + 4$, $2 + 3$, $3 + 2$, or $4 + 1$. Hence the probability that the sum of the dice is 5 is $\frac{4}{36} = \frac{1}{9}$. The sum can be 12 in only one way, $6 + 6$. Hence the probability that the sum is 12 is $\frac{1}{36}$. $\diamond$

Example 4.5 Suppose that two cards are drawn successively from a deck of cards. What is the probability that both are hearts? There are 52 possibilities for the first card, and for each of these there are 51 possibilities for the second. Hence there are $52 \cdot 51$ possibilities for the result of the two draws. We assign the equiprobable measure. The statement "The two cards are hearts" is true in $13 \cdot 12$ of the $52 \cdot 51$

possibilities. Hence the probability of this statement is $13 \cdot 12/(52 \cdot 51) = \frac{1}{17}$. ◇

Example 4.6 Assume that, on the basis of a predictive index applied to students A, B, and C when entering college, it is predicted that after four years of college the scholastic record of A will be the highest, C the second highest, and B the lowest of the three. Suppose, in fact, that these predictions turn out to be exactly correct. If the predictive index has no merit at all and hence the predictions amount simply to guessing, what is the probability that such a prediction will be correct? There are $3! = 6$ orders in which the men might finish. If the predictions were really just guessing, then we would assign an equal weight to each of the six outcomes. In this case the probability that a particular prediction is true is $\frac{1}{6}$. Since this probability is reasonably large, we would hesitate to conclude that the predictive index is in fact useful, on the basis of this one experiment. Suppose, on the other hand, it predicted the order of six men correctly. Then a similar analysis would show that, by guessing, the probability is $\frac{1}{6!} = \frac{1}{720}$ that such a prediction would be correct. Hence, we might conclude here that there is strong evidence that the index has some merit. ◇

Exercises

1. A letter is chosen at random from the word "random". What is the probability that it is an n? That it is a vowel?

   [Ans. $\frac{1}{6}$;$\frac{1}{3}$.]

2. An integer between 3 and 12 inclusive is chosen at random. What is the probability that it is an even number? That it is even and divisible by three?

3. A card is drawn at random from a pack of playing cards.

   (a) What is the probability that it is either a heart or the king of clubs?

   [Ans. $\frac{7}{26}$.]

   (b) What is the probability that it is either the queen of hearts or an honor card (i.e., ten, jack, queen, king, or ace)?

[Ans. $\frac{5}{13}$.]

4. A word is chosen at random from the set of words

$$\mathcal{U} = \{\text{men}, \text{bird}, \text{ball}, \text{field}, \text{book}\}.$$

Let p, q, and r be the statements:

p: The word has two vowels.

q: The first letter of the word is b.

r: The word rhymes with cook.

Find the probability of the following statements.

(a) p.

(b) q.

(c) r.

(d) $p \wedge q$.

(e) $(p \vee q) \wedge \neg r$.

(f) $p \rightarrow q$.

[Ans. $\frac{4}{5}$.]

5. A single die is thrown. Find the probability that

(a) An odd number turns up.

(b) The number which turns up is greater than two.

(c) A seven turns up.

6. In the Primary voting example of Section 2.1, assume that all 36 possibilities in the elections are equally likely. Find

(a) The probability that candidate A wins more states than either B or C.

[Ans. $\frac{7}{18}$.]

(b) That all the states are won by the same candidate.

[Ans. $\frac{1}{36}$.]

(c) That every state is won by a different candidate.

[Ans. 0.]

7. A single die is thrown twice. What value for the sum of the two outcomes has the highest probability? What value or values of the sum has the lowest probability of occurring?

8. Two boys and two girls are placed at random in a row for a picture. What is the probability that the boys and girls alternate in the picture?

[Ans. $\frac{1}{3}$.]

9. A certain college has 500 students and it is known that

300	read French.
200	read German.
50	read Russian.
20	read French and Russian.
30	read German and Russian.
20	read German and French.
10	read all three languages.

If a student is chosen at random from the school, what is the probability that the student

(a) Reads two and only two languages?

(b) Reads at least one language?

10. Suppose that three people enter a restaurant which has a row of six seats. If they choose their seats at random, what is the probability that they sit with no seats between them? What is the probability that there is at least one empty seat between any two of them?

11. Find the probability of obtaining each of the following poker hands. (A poker hand is a set of five cards chosen at random from a deck of 52 cards.)

(a) Royal flush (ten, jack, queen, king, ace in a single suit).

[Ans. $4/\binom{52}{5} = .0000015$.]

(b) Straight flush (five in a sequence in a single suit, but not a royal flush).

[Ans. $(40 - 4)/\binom{52}{5} = .000014.$]

(c) Four of a kind (four cards of the same face value).

[Ans. $624/\binom{52}{5} = .00024.$]

(d) Full house (one pair and one triple of the same face value).

[Ans. $3744/\binom{52}{5} = .0014.$]

(e) Flush (five cards in a single suit but not a straight or royal flush).

[Ans. $(5148 - 40)/\binom{52}{5} = .0020.$]

(f) Straight (five cards in a row, not all of the same suit).

[Ans. $(10,240 - 40)/\binom{52}{5} = .0039.$]

(g) Straight or better.

[Ans. .0076.]

12. If ten people are seated at a circular table at random, what is the probability that a particular pair of people are seated next to each other?

[Ans. $\frac{2}{9}$.]

13. A room contains a group of n people who are wearing badges numbered from 1 to n. If two people are selected at random, what is the probability that the larger badge number is a 3? Answer this problem assuming that $n = 5, 4, 3, 2$.

[Ans. $\frac{1}{5}; \frac{1}{3}; \frac{2}{3}; 0.$]

14. In Exercise 13, suppose that we observe two men leaving the room and that the larger of their badge numbers is 3. What might we guess as to the number of people in the room?

15. Find the probability that a bridge hand will have suits of

(a) 5, 4, 3, and 1 cards.

[Ans. $\frac{4!\binom{13}{5}\binom{13}{4}\binom{13}{3}\binom{13}{1}}{\binom{52}{13}} = .129.$]

(b) 6, 4, 2, and 1 cards.

[Ans. .047.]

(c) 4, 4, 3, and 2 cards.

[Ans. .216.]

(d) 4, 3, 3, and 3 cards.

[Ans. .105.]

16. There are $\binom{52}{13} = 6.5 \times 10^{11}$ possible bridge hands. Find the probability that a bridge hand dealt at random will be all of one suit. Estimate roughly the number of bridge hands dealt in the entire country in a year. Is it likely that a hand of all one suit will occur sometime during the year in the United States?

Supplementary exercises.

17. Find the probability of not having a pair in a hand of poker.

18. Find the probability of a "bust" hand in poker. [Hint: A hand is a "bust" if there is no pair, and it is neither a straight nor a flush.]

[Ans. .5012.]

19. In poker, find the probability of having

(a) Exactly one pair.

[Ans. .4226.]

(b) Two pairs.

[Ans. .0475.]

(c) Three of a kind.

[Ans. .0211.]

20. Verify from Exercises 11, 18, 19 that the probabilities for all possible poker hands add up to one (within a rounding error).

21. A certain French professor announces that he or she will select three out of eight pages of text to put on an examination and that each student can choose one of these three pages to translate.

(a) What is the maximum number of pages that a student should prepare in order to be certain of being able to translate a page that he or she has studied?

(b) Smith decides to study only four of the eight pages. What is the probability that one of these four pages will appear on the examination?

Two nonintuitive examples

There are occasions in probability theory when one finds a problem for which the answer, based on probability theory, is not at all in agreement with one's intuition. It is usually possible to arrange a few wagers that will bring one's intuition into line with the mathematical theory. A particularly good example of this is provided by the matching birthdays problem.

Assume that we have a room with r people in it and we propose the bet that there are at least two people in the room having the same birthday, i.e., the same month and day of the year. We ask for the value of r which will make this a fair bet. Few people would be willing to bet even money on this wager unless there were at least 100 people in the room. Most people would suggest 150 as a reasonable number. However, we shall see that with 150 people the odds are approximately 4,100,000,000,000,000 to 1 in favor of two people having the same birthday, and that one should be willing to bet even money with as few as 23 people in the room.

Let us first find the probability that in a room with r people, no two have the same birthday. There are 365 possibilities for each person's birthday (neglecting February 29). There are then 365^r possibilities for the birthdays of r people. We assume that all these possibilities are equally likely. To find the probability that no two have the same birthday we must find the number of possibilities for the birthdays which have no day represented twice. The first person can have any of 365 days for his or her birthday. For each of these, if the second person is to have a different birthday, there are only 364 possibilities for his or her birthday. For the third person, there are 363 possibilities if he or she is to have a different birthday than the first two, etc. Thus the probability that no two people have the same birthday in a group of r people is

$$q_r = \frac{365 \cdot 364 \cdot \ldots \cdot (365 - r + 1)}{365^r}.$$

Number of people in the room	Probability of at least two with same birthday	Approximate odds for a fair bet
5	.027	
10	.117	
15	.253	
20	.411	70:100
21	.444	80:100
22	.476	91:100
23	.507	103:100
24	.538	117:100
25	.569	132:100
30	.706	241:100
40	.891	819:100
50	.970	33:1
60	.994	170:1
70		1,200:1
80		12,000:1
90		160,000:1
100		3,300,000:1
125		31,000,000,000:1
150		4,100,000,000,000,000:1

Figure 4.1: ♢

The probability that at least two people have the same birthday is then $p_r = 1 - q_r$. In Figure 4.1 the values of p_r and the odds for a fair bet, $p_r : (1 - p_r)$ are given for several values of r.

We consider now a second problem in which intuition does not lead to the correct answer. A hat-check clerk has checked n hats, but they have become hopelessly scrambled. The clerk hands back the hats at random. What is the probability that at least one head gets its own hat? For this problem some people's intuition would lead them to guess that for a large number of hats this probability should be small, while others guess that it should be large. Few people guess that the probability is neither large nor small and essentially independent of the number of hats involved.

Let p_j be the statement "the jth head gets its own hat back". We wish to find $\Pr[p_1 \vee p_2 \vee \ldots \vee p_n]$. We know from Exercise 26 that a probability of this form can be found from the inclusion-exclusion formula. We must add all probabilities of the form $\Pr[p_i]$, then subtract the sum of all probabilities of the form $\Pr[p_i \wedge p_j]$, then add the sum of all probabilities of the form $\Pr[p_i \wedge p_j \wedge p_k]$, etc.

However, each of these probabilities represents the probability that a particular set of heads get their own hats back. These probabilities are very easy to compute. Let us find the probability that out of n heads some particular m of them get back their own hats. There are $n!$ ways that the hats can be returned. If a particular m of them are to get their own hats there are only $(n-m)!$ ways that it can be done. Hence the probability that a particular m heads get their own hats back is

$$\frac{(n-m)!}{n!}.$$

There are $\binom{n}{m}$ different ways we can choose m heads out of n. Hence the mth group of terms contributes

$$\binom{n}{m} \cdot \frac{(n-m)!}{n!} = \frac{1}{m!}$$

to the alternating sum. Thus

$$\Pr[p_1 \vee p_2 \vee \ldots \vee p_n] = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots \pm \frac{1}{n!},$$

where the $+$ sign is chosen if n is odd and the $-$ sign if n is even. In Figure 4.2, these numbers are given for the first few values of n.

It can be shown that, as the number of hats increases, the probabilities approach a number $1 - (1/e) = .632121\ldots$, where the number $e = 2.71828\ldots$ is a number that plays an important role in many branches of mathematics.

Exercises

1. What odds should you be willing to give on a bet that at least two people in the United States Senate have the same birthday?

[Ans. $3,300,000 : 1$.]

Number of hats	Probability p_n that at least one man gets his own hat
2	.500000
3	.666667
4	.625000
5	.633333
6	.631944
7	.632143
8	.632118

Figure 4.2: ◇

2. What is the probability that in the House of Representatives at least two men have the same birthday?

3. What odds should you be willing to give on a bet that at at least two of the Presidents of the United States have had the same birthday? Would you win the bet?

   [Ans. More than 4 : 1; Yes. Polk and Harding were born on Nov. 2.]

4. What odds should you be willing to give on the bet that at least two of the Presidents of the United States have died on the same day of the year? Would you win the bet?

   [Ans. More than 2.7 : 1; Yes. Jefferson, Adams, and Monroe all died on July 4.]

5. Four men check their hats. Assuming that the hats are returned at random, what is the probability that *exactly* four men get their own hats? Calculate the answer for 3, 2, 1, 0 men.

   [Ans. $\frac{1}{24}$; 0; $\frac{1}{4}$; $\frac{1}{3}$; $\frac{3}{8}$.]

6. A group of 50 knives and forks a dance. The partners for a dance are chosen by lot (knives dance with forks). What is the approximate probability that no knife dances with its own fork?

Supplementary exercises.

Find a formula for the probability of having more than one coincidence of birthdays among n people, i.e., of having at least two pairs of identical birthdays, or of three or more people having the same birthday. [Hint: Take the probability of at least one coincidence, and subtract the probability of having exactly one pair.]

Compute the probability of having more than one coincidence of birthdays when there are 20, 25, 30, 40, or 50 people in the room.

What is the smallest number of people you need in order to have a better than even chance of finding more than one coincidence of birthdays?

[Ans. 36.]