General Mathematics and Computational Science I

Combinatorics and Probability

Marcel Oliver

Abstract

Problems in probability theory often hinge on combinatorics. This note will introduce the notion of empirical probability as the relative frequency of the occurrence of certain events—its computation thus requires careful and often nontrivial counting. We will illustrate the concept with elementary examples. This note is not meant to give a complete and thorough overview of discrete probability theory.

Problem 1. Players A and B each toss a die. Player B wins if he scores higher than A, otherwise he loses. Is this a fair game?

Let's count the the number of possible pairs of scores where player B wins. If A throws a ⚅, B loses. If A throws a ⚄, there is exactly one possibility for B to win; if B throws a ⚃, there are 2 possibilities for B to win, and so on. Thus, the total number of ways for B to win is

$$0 + 1 + 2 + 3 + 4 + 5 = 15\,. \tag{1}$$

We could count the number of possibilities for A to win the same way; the following argument, however, is simpler. There are $6 \cdot 6 = 36$ possible ways the two (ordered!) dice can fall. Thus, the total number of ways for A to win is

$$36 - 15 = 21\,. \tag{2}$$

We conclude that the game is $\frac{21}{15} = \frac{7}{5}$ in A's favor; the game is not fair.

In the language of probability theory, the above can be formalized as follows. We speak of a *trial* each time an experiment, such as the tossing of dice, is performed. A possible result of a trial is called an *outcome*. It is always assumed that trials are *independent*, i.e. that the outcome of one trial does not depend on any past trials. The set of all possible outcomes is called the *sample space*, usually denoted S. Outcomes can be grouped into *events*, which formally are subsets $A \subseteq S$. Finally, the *probability* of an event A occurring is the relative frequency of that event in a large number of trials. Assuming that every outcome is equally likely, it is given by

$$P(A) = \frac{n_A}{n_S} = \frac{\text{number of outcomes in } A}{\text{total number of possible outcomes}}\,. \tag{3}$$

Example 1. In Problem 1, the sample space is the set of ordered tuples of scores,

$$S = \{\{\text{⚀⚀}\}, \{\text{⚀⚁}\}, \ldots, \{\text{⚀⚅}\}, \{\text{⚁⚀}\}, \ldots, \{\text{⚅⚅}\}\} . \tag{4}$$

We assign the first die in each tuple to player A, the second to player B. Then the event "player A wins" is the set

$$A = \{\{\text{⚀⚀}\}, \{\text{⚁⚀}\}, \{\text{⚁⚁}\}, \{\text{⚂⚀}\}, \{\text{⚂⚁}\}, \{\text{⚂⚂}\}, \ldots, \{\text{⚅⚅}\}\} , \tag{5}$$

so that

$$P(A) = \frac{n_A}{n_S} = \frac{21}{36} = \frac{7}{12} . \tag{6}$$

In the game of Yahtzee (or Yacht), five dice are tossed. Players score by obtaining special patterns, such as the *Yahtzee* (five of the same kind, e.g. ⚄⚄⚄⚄⚄), *full house* (two and three of the same kind but not a Yahtzee; e.g. ⚁⚁⚃⚃⚃), *small straight* (four in a row, e.g. ⚁⚁⚂⚃⚄) or *large straight* (five in a row, e.g. ⚀⚁⚂⚃⚄). The actual rules of the game allow players to twice repeat tossing a select subset of dice. In the following, however, we disregard the option to repeat; for details on the mathematics of this game, see [2].

Problem 2. Find the probability of throwing a Yahtzee.

The total number of possible outcomes is 6^5, as each of the five dice can independently score one of six values. Note that in this count the dice are ordered, whereas the scoring rules disregard order. The number of possible Yahtzees is 6, one for each of the scores ⚀, …, ⚅. We conclude that that

$$P(\text{Yahtzee}) = \frac{6}{6^5} = \frac{1}{6^4} = \frac{1}{1296} . \tag{7}$$

Problem 3. Find the probability of throwing a full house.

The set of unordered full houses has 6 (for the first number kind) times 5 (for the second number kind which must be different from the first) members. Since our sample space consists of *ordered* 5-tuples, we must multiply this number by the number of ways a full house can be realized with five ordered dice, i.e., by the number of 5-term binary sequences containing three 1s and two 0s. We obtain

$$P(\text{full house}) = \frac{5 \cdot 6}{6^5} \frac{5!}{2!\,3!} = \frac{25}{648} . \tag{8}$$

Problem 4. What is the probability that within a group of n people at least two have the same birthday?

It is often easier to count the number of possibilities in which an event does *not* occur; this is the case here. Ignoring leap years, the total number of ways n ordered people can have their birthdays is 365^n. The number of ways these birthdays can fall without coinciding is counted as follows. The first person can be born on any of the 365 days of the year, the second on any of the remaining 364, the third on any of the remaining 363, and so on. We find that

$$P(\text{birthdays all different}) = \frac{365 \cdot 364 \cdot \dots \cdot (365 - n + 1)}{365^n} = \frac{365!}{365^n\,(365 - n)!}\,, \tag{9}$$

and therefore

$$P(\text{at least two birthdays coincide}) = 1 - P(\text{birthdays all different})\,. \tag{10}$$

For our class of 20 students, this expression evaluates to approximately 0.41.

Problem 5. *(Hat check problem.)* 12 men leave their hat with the hat check. They lose their tickets so that the hats get returned at random. What is the probability that nobody obtains his own hat?

Consider the problem for n hats. If w_n denotes that number of ways that n hats can be returned to n men without any getting his own (such permutations are called *derangements*), a careful count (how?) leads to the two-term recurrence

$$w_{n+1} = n\,w_n + n\,w_{n-1}\,, \tag{11}$$

which implies that

$$w_n = n\,w_{n-1} + (-1)^n\,. \tag{12}$$

The one-term recurrence (12) can be proved by induction (what are the starting values w_1 and w_2?); there is also a very elegant derivation using generating functions [3] which, however, goes far beyond what we discussed in this class. Letting $p_n = w_n/n!$ denote the associated probabilities, we conclude that

$$p_n = \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!}\,. \tag{13}$$

A detailed discussion of this problem, and much more, can be found in [1].

References

[1] C.M. Grinstead and J.L. Snell, *Introduction to Probability*, AMS, 1997.
http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/book.html

[2] E.W. Weisstein, *Yahtzee*. From MathWorld, Wolfram Research.
http://mathworld.wolfram.com/Yahtzee.html

[3] H.S. Wilf, *Generatingfunctionology*, Academic Press, 1994.
http://www.math.upenn.edu/~wilf/DownldGF.html