Reprinted from the AMERICAN MATHEMATICAL MONTHLY
Vol. 71, No. 7, August-September, 1964

REPEATED INDEPENDENT TRIALS AND A CLASS OF DICE PROBLEMS

EDWARD O. THORP, New Mexico State University

1. Introduction. This discussion was originally motivated by a class of dice problems. They are illustrated by the following examples, which will be referred to in the sequel.

Assume that two true dice are rolled repeatedly.

Problem 1. Find the probability that both the totals 5 and 9 appear before a 7 appears.

Problem 2. Find the probability that both the totals 4 and 6 appear before a 7 appears.

Problem 3. Find the probability that 4, 5, 6, 8, 9, 10 all appear before 7 appears.

Problem 4. Find the probability that all totals different from 7 appear before a 7 appears.

Problem 1 is intuitively quite simple when we observe that on any one trial $P(5)=P(9)=4/36$ and $P(7)=6/36$, where $P(T)$ is the probability that a total of T occurs on any given trial. We might argue loosely that the probability that either a 5 or a 9 occurs before a 7 is $8/14$. The probability that the other one then occurs before a 7 is $4/10$. The probability that both 5 and 9 appear before 7 is thus $(8/14)(4/10)=8/35$.

Problem 2 is surprisingly difficult by comparison with problem 1. This is due to the fact that $P(4)=3/36\neq 5/36=P(6)$. We give below the solution to a general problem concerning repeated independent trials, of which problems 2, 3 and 4 are special cases which we will solve as illustrations. Finally we discuss some useful approximations to the general solution.

2. Formal solution of a class of problems. Consider a series of repeated independent trials with the outcomes of each trial being events in a given (fixed) sample space. Let $E_1, \cdots, E_m, B$, be $m+1$ events with B disjoint from each of the E_i. What is the probability that all the events E_i will occur before the event B occurs?

Let A_1^i ($A_2^i, \cdots, A_m^i, B^i$, respectively) be the event that E_1 ($E_2, \cdots, E_m, B$, respectively) does not occur in the first i trials. Let B_{i+1} be the event "B occurs on trial $i+1$" and let F_{i+1} be the event $(A_1^i \cup \cdots \cup A_m^i)B^iB_{i+1}$,where $i=1, 2, \cdots$. Let $F_1=B_1$. Thus F_i is the event that B occurs for the first time on trial i and not all the $E_1, \cdots, E_m$ have yet occurred. Hereafter we refer to

F_i as failure at the ith trial. The probability $P(F)$ that B occurs before all the E_i, which event we refer to as failure, is clearly $\sum_{i=0}^{\infty} P(F_{i+1})$. The solution to our problem is then $1-P(F)$.

Setting $A_1^i \cup \cdots \cup A_n^i = A^i$, we see that

$$P(F_{i+1}) = P(A^iB^iB_{i+1}) = P(A^iB^i)P(B_{i+1}) = P(A^iB^i)P(B),$$

where the second equality follows because an event occurring on trial $i+1$ is stochastically independent of all events, such as A^iB^i, occurring on the first i trials. It remains to calculate $P(A^iB^i)$. For $i \geqq 1$,

$$P(A^iB^i) = P\left(\bigcup_{k=1}^{m} A_k^iB^i\right) = \sum_{k_1} P(A_{k_1}^iB^i) - \sum_{k_1<k_2\le m} P(A_{k_1}^iA_{k_2}^iB^i) + \cdots$$
$$= \sum_{k_1} (P(A_{k_1}B^1))^i - \sum_{k_1<k_2\le m} (P(A_{k_1}A_{k_2}B^1))^i + \cdots$$

where the second equality follows from a well-known formula (see, e.g. [1] p. 89) and the third equality follows from the independence of trials.

The probability of eventual failure is given by

$$P(F) = \sum_{i=0}^{\infty} P(F_{i+1}) = P(B)\left\{1 + \sum_{i=1}^{\infty} P(A^iB^i)\right\}$$
$$= P(B)\left\{\sum_{k_1}\sum_{i=0}^{\infty} (P(A_{k_1}B^1))^i - \sum_{k_1<k_2\le m}\sum_{i=0}^{\infty} (P(A_{k_1}A_{k_2}B^1))^i + \cdots\right\}$$
$$= P(B)\left\{\sum_{k_1} \frac{1}{1-P(A_{k_1}B^1)} - \sum_{k_1<k_2\le m} \frac{1}{1-P(A_{k_1}A_{k_2}B^1)} + \cdots\right\}$$
$$= P(B)\left\{\sum_{k_1} \frac{1}{P(E_{k_1}\cup B)} - \sum_{k_1<k_2\le m} \frac{1}{P(E_{k_1}\cup E_{k_2}\cup B)} + \cdots\right\}.$$

The probability of eventual success is therefore

$$P(S) = P(B)\left\{\frac{1}{P(B)} - \sum_{k_1} \frac{1}{P(E_{k_1}\cup B)} + \cdots\right\} \qquad \text{when } P(B) > 0.$$

The third equality in the calculation for $P(F)$ follows, i.e., the sums can be taken from $i=0$, because

$$\sum_{k_1}(1) - \sum_{k_1<k_2\le m}(1) + \cdots = 1 - (1-1)^m = 1.$$

When $P(B)=0$, it can be shown that $P(S)=1$ if all $P(E_i)>0$ and $P(S)=0$ if some $P(E_i)=0$.

Applying this to problem 2, we have $P(F)=(6/36)\{36/9+36/11-36/14\}$ $=181/231$ whence $P(S)$, the probability of success, is $50/231$. In problem 3 the probability of success is .062164; the odds against are 15.09:1. In the solution of problem 3, however, $2^6-1=63$ terms appear and, even after using the symmetries in the problem (i.e., that $P(3)=P(10)$, $P(4)=P(9)$, $P(5)=P(8)$), some 24 terms need to be considered. In the solution to problem 4, $2^{10}-1=1023$ terms appear, but the form of the solution makes machine calculation quite easy. If it is necessary to solve the problem by hand, however, then, despite symmetries, 210 terms need to be considered. Even though the solution reduces to an expression of the form $\sum_{n=7}^{36} p_n/n$, where the p_n are integers, the calculation is tedious. In the general case even these simplifications do not occur. Approximations to the solution are desirable.

3. Numerical solutions to problems. For disjoint events it is possible to show, by using an argument based on a tree diagram, that

$$P(S) = P(E_1)\cdots P(E_m)\sum_s 1/\{[P(B)+P(E_{s(1)})]\cdots[P(B)+P(E_{s(1)})+\cdots + P(E_{s(m)})]\},$$

where the sum is over all permutations s of the integers from 1 to m.

When $P(E_1)=\cdots=P(E_m)=k/n$ and $P(B)=b/n$, it is fairly easy to verify directly that the two forms for $P(S)$ are identical. Since

$$\binom{m}{1}-\binom{m}{2}+\cdots = 1-(1-1)^m = 1,$$

we see that

$$\begin{aligned}\frac{k}{k+b}\cdots\frac{(m-1)k}{(m-1)k+b}\cdot\frac{mk}{mk+b} &\overset{?}{=} b\left\{\frac{1}{b}-\binom{m}{1}\frac{1}{k+b}+\binom{m}{2}\frac{1}{2k+b}-\cdots\right\}\\ &= b\left\{\binom{m}{1}\left(\frac{1}{b}-\frac{1}{k+b}\right)-\binom{m}{2}\left(\frac{1}{b}-\frac{1}{2k+b}\right)+\cdots\right\}\\ &= \left\{\binom{m}{1}\frac{k}{k+b}-\binom{m}{2}\frac{2k}{2k+b}+\cdots\right\}.\end{aligned}$$

Letting $b/k=x$, our verification reduces to

$$\frac{1}{1+x}\cdots\frac{(m-1)}{(m-1)+x}\cdot\frac{m}{m+x}\overset{?}{=}\binom{m}{1}\frac{1}{1+x}-\binom{m}{2}\frac{2}{2+x}+\cdots.$$

When we compute the partial fractions expansion of the left side, we find that it indeed coincides with the right hand expression. The direct verification that the two forms for $P(S)$ agree in the general case does not seem as easy.

We conjecture that, in the case of disjoint events, the inequalities

$$L=\frac{m!P(E_1)\cdots P(E_m)}{[\mu+P(E)]\cdots[m\mu+P(B)]}\leqq P(S)\leqq\frac{m!\mu^m}{[\mu+P(B)]\cdots[m\mu+P(B)]}=U$$

are satisfied. We expected the alternate form for $P(S)$, given above, to be useful. We have been unable, however, either to prove or to disprove the conjecture.

The conjecture is true for $m=1$ (trivially) and for $m=2$ (where it reduces to the fact that $4ab \leqq (a+b)^2$). Also, when $P(E_1) = \cdots = P(E_m)$, the two conjectured bounds coincide and the resulting number agrees with the alternate form for $P(S)$. The conjecture holds for $P(B)=0$ or 1.

Using an electronic computer and the first form of the solution for $P(S)$, we solved problems 3 and 4. The results, shown in the Table, are again in agreement with the conjecture.

TABLE 1 (Computed by R. L. South)

Problem	$P(F)$	L	$P(S)$	U	*True odds to 1 against S*	*U odds to 1 against S*
1	0.7714	0.2286	0.2286	0.2286	3.375	3.375
2	0.7835	0.2143	0.2165	0.2286	3.620	3.375
3	0.9378	0.0599	0.0622	0.0682	15.085	13.663
3′	0.9926	0.0054	0.0074	0.0222	134.135	44.000
4	0.99474	0.00369	0.00526	0.01515	189.20	65.000

Problem 3′ is to find the probability that 2, 3, 5, 6, 8, 9, 11, 12 all appear before 7. Problems 3 and 4 first came to my attention from [2] page 665, where they are solved incorrectly via U.

References

1. W. Feller, An Introduction to Probability Theory and Its Applications, Vol. I, 2nd ed. Wiley, New York, 1957.
2. John Scarne, Scarne's Complete Guide to Gambling, Simon and Shuster, New York, 1961.