

**EX:** Five fair 6-sided dice are rolled. Find the probability that three or more dice show the same number.

**SOL'N:** In this problem, difficulty arises from the overlap of rolls containing, for example, four of the same number and rolls containing three of the same number. A roll containing four of the same number also contains three of the same number. When identifying outcomes, we must avoid over-counting.

The solution to this dilemma is to find a convenient partition of the sample space that includes events that consist of exactly the outcomes of interest, namely rolls where three or more dice show the same number. The term "convenient" means the probability of each event in the partition is easily calculated. By definition, the events in a partition are non-overlapping but include all possible outcomes, (like the pieces of glass that form a complete stained-glass window). Also, the probability of a union of events in a partition is equal to the sum of the probabilities of those events. Thus, our final answer will be a sum of probabilities.

Here, a convenient partition involves the maximum number of dice showing the same number.

$A_1 \equiv$  All five dice show the same number

$A_2 \equiv$  Exactly four dice show the same number (the other die must not be the same number as the four-of-a-kind)

$A_3 \equiv$  Exactly three dice show the same number (the other two may form a pair but must not be the same number as the three-of-a-kind)

$A_4 \equiv$  No more than one or two dice show the same number

Our answer will be the sum of probabilities of  $A_1$  through  $A_3$ , (we define  $A_4$  merely to complete the partition).

$$P(\text{three or more dice show the same number}) = P(A_1) + P(A_2) + P(A_3)$$

**NOTE:** We must be certain that our partition includes all possible outcomes and consists of non-overlapping events. Short of listing every possible outcome and verifying that each one falls into one and only one of the events in the partition, it is difficult to rigorously prove we have a partition. We must instead rely on experience, careful thought, common sense, and inspection of examples to convince ourselves that our partition is valid. While the same might be said of mathematical proofs in general, the complexity of possible outcomes makes combinatorics a particularly daunting subject. Consider, for example, the probability that two and no more than two dice show the same number. The possibility of having two pair complicates the calculation significantly.

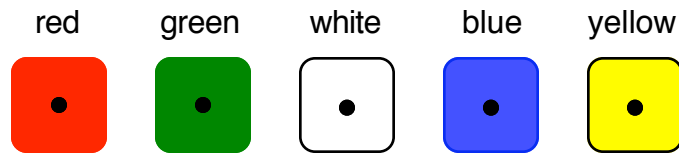
**NOTE:** We may lump all outcomes outside the set of interest into one event defined as "all other outcomes". Here, that event is  $A_4$ . Giving a more precise definition of  $A_4$  serves the purpose of clarifying that we have no outcomes in  $A_4$  that we should have included in  $A_1$  through  $A_3$ .

Given that the dice are fair, the probabilities for all possible outcomes are equal. Since there are five six-sided dice, the number of possible outcomes is  $6^5$ . It follows that the probability of an event is the number of outcomes in that event divided by  $6^5$ .

The number of outcomes in  $A_1$ , all five dice showing the same number, is the number of choices for that number. Since there are six numbers on the dice, the number of outcome in  $A_1$  is 6.

$$P(A_1) = \frac{6}{6^5}$$

Another way to view this is to imagine that the dice are different colors and rolled one at a time in a color order, as shown below. This yields the same probabilities as rolling five identical dice all at once.



There are 6 choices for the number on the first die. After that, all four remaining dice must have 1 particular number. Thus, the number of outcomes in  $A_1$  is  $6 \cdot 1 \cdot 1 \cdot 1 \cdot 1 = 6$ , as argued above.

A third way to compute the number of outcomes in  $A_1$  is to pick the number showing on the dice and then determine how many ways there are to choose which five dice have that number showing. There are 6 choices for the number showing on the dice, and there is  ${}_5C_5 = 1$  way to choose which 5 dice show that number. Thus, there are  $6 \cdot 1 = 6$  outcomes for five-of-a-kind.

Proceeding in the same fashion, the number of outcomes in  $A_2$  is the product of the following terms:

1. The number of choices for the number showing on the dice for the four-of-a-kind = 6
2. The number of ways of choosing which four dice the four-of-a-kind appear on =  ${}_5C_4 = \frac{5!}{1!4!} = 5$
3. The number of choices for the number showing on the die that is not in the four-of-a-kind =  $6 - 1 = 5$ , since the number must be different from the four-of-a-kind.
4. The number of ways of choosing which die is not part of the four-of-a-kind =  ${}_1C_1 = 1$ . Note that, because the four-of-a-kind are removed from the original set of five dice before we choose which die is not part of the four-of-a-kind, we are picking from a set of one object.

Thus, the number of outcomes in  $A_2$  is  $6 \cdot 5 \cdot 5 \cdot 1 = 150$ .

$$P(A_2) = \frac{150}{6^5}$$

The number of outcomes in  $A_3$  is the product of the following terms:

1. The number of choices for the number showing on the dice for the three-of-a-kind = 6
2. The number of ways of choosing which three dice the three-of-a-kind appear on =  ${}_5C_3 = \frac{5!}{2!3!} = 10$
3. The number of choices for the numbers showing on the dice that are not in the three-of-a-kind =  $(6 - 1)^2 = 5^2 = 25$ , since the number must be different from the three-of-a-kind. Note that we could have a pair on the remaining two dice and, if we were trying to count pairs, this would complicate matters significantly. Here, however, a pair is acceptable since this is fewer than three-of-a-kind.
4. The number of ways of choosing which die is not part of the three-of-a-kind =  ${}_2C_2 = 1$ . Note that, because the three-of-a-kind are removed from the original set of five dice before we choose which dice are not part of the three-of-a-kind, we are picking from a set of two objects.

Thus, the number of outcomes in  $A_3$  is  $6 \cdot 10 \cdot 25 \cdot 1 = 1500$ .

$$P(A_3) = \frac{1500}{6^5}$$

Summing the probabilities for  $A_1$ ,  $A_2$ , and  $A_3$  gives our final answer:

$$P(3 \text{ or more dice show same number}) = \frac{6 + 150 + 1500}{6^5} = 0.213$$

This is about 1 in 5.