Video

2.1.2 Ordered Sampling without Replacement: Permutations

Consider the same setting as above, but now repetition is not allowed. For example, if $A = \{1, 2, 3\}$ and $k = 2$, there are 6 different possibilities:

1. (1,2);
2. (1,3);
3. (2,1);
4. (2,3);
5. (3,1);
6. (3,2).

In general, we can argue that there are k positions in the chosen list: $($Position 1, Position 2, ..., Position $k)$. There are n options for the first position, $(n-1)$ options for the second position (since one element has already been allocated to the first position and cannot be chosen here), $(n-2)$ options for the third position, ... $(n-k+1)$ options for the kth position. Thus, when ordering matters and repetition is not allowed, the total number of ways to choose k objects from a set with n elements is

$$n \times (n-1) \times \ldots \times (n-k+1).$$

Any of the chosen lists in the above setting (choose k elements, ordered and no repetition) is called a k-permutation of the elements in set A. We use the following notation to show the number of k-permutations of an n-element set:

$$P^n_k = n \times (n-1) \times \ldots \times (n-k+1).$$

Note that if k is larger than n, then $P^n_k = 0$. This makes sense, since if $k > n$ there is no way to choose k distinct elements from an n-element set. Let's look at a very famous problem, called the birthday problem, or the birthday paradox.

Example 2.4

If k people are at a party, what is the probability that at least two of them have the same birthday? Suppose that there are $n = 365$ days in a year and all days are equally likely to be the birthday of a specific person.

Solution

Let A be the event that at least two people have the same birthday. First note that if $k > n$, then $P(A) = 1$; so, let's focus on the more interesting case where $k \leq n$. Again, the phrase "at least" suggests that it might be easier to find the probability of the complement event, $P(A^c)$. This is the event that no two people have the same birthday, and we have

$$P(A) = 1 - \frac{|A^c|}{|S|}.$$

Thus, to solve the problem it suffices to find $|A^c|$ and $|S|$. Let's first find $|S|$. What is the total number of possible sequences of birthdays of k people? Well, there are $n = 365$ choices for the first person, $n = 365$ choices for the second person,... $n = 365$ choices for the kth person. Thus there are

$$n^k$$

possibilities. This is, in fact, an ordered sampling with replacement problem, and as we have discussed, the answer should be n^k (here we draw k samples, birthdays, from the set $\{1, 2, \ldots, n = 365\}$). Now let's find $|A^c|$. If no birthdays are the same, this is similar to finding $|S|$ with the difference that repetition is not allowed, so we have

$$|A^c| = P^n_k = n \times (n-1) \times \ldots \times (n-k+1).$$

You can see this directly by noting that there are $n = 365$ choices for the first person, $n - 1 = 364$ choices for the second person,..., $n - k + 1$ choices for the kth person. Thus the probability of A can be found as

$$\begin{aligned} P(A) &= 1 - \frac{|A^c|}{|S|} \\ &= 1 - \frac{P^n_k}{n^k}. \end{aligned}$$

Discussion: The reason this is called a paradox is that $P(A)$ is numerically different from what most people expect. For example, if there are k=23 people in the party, what do you guess is the probability that at least two of them have the same birthday, $P(A)$? The answer is $.5073$, which is much higher than what most people guess. The probability crosses 99 percent when the number of peoples reaches 57. But why is the probability higher than what we expect?

It is important to note that in the birthday problem, neither of the two people are chosen beforehand. To better answer this question, let us look at a different problem: I am in a party with $k-1$ people. What is the probability that at least one person in the party has the same birthday as mine? Well, we need to choose the birthdays of $k-1$ people, the total number of ways to do this is n^{k-1}. The total number of ways to choose the birthdays so that no one has my birthday is $(n-1)^{k-1}$. Thus, the probability that at least one person has the same birthday as mine is

$$P(B) = 1 - (\frac{n-1}{n})^{k-1}.$$

Now, if $k = 23$, this probability is only $P(B) = 0.0586$, which is much smaller than the corresponding $P(A) = 0.5073$. The reason is that event B is looking only at the case where one person in the party has the same birthday as me. This is a much smaller event than event A which looks at all possible pairs of people. Thus, $P(A)$ is much larger than $P(B)$. We might guess that the value of $P(A)$ is much lower than it actually is, because we might confuse it with $P(B)$.

Permutations of n elements: An n-permutation of n elements is just called a permutation of those elements. In this case, $k = n$ and we have

$$\begin{aligned} P_n^n &= n \times (n-1) \times \ldots \times (n-n+1) \\ &= n \times (n-1) \times \ldots \times 1, \end{aligned}$$

which is denoted by $n!$, pronounced "n factorial". Thus $n!$ is simply the total number of permutations of n elements, i.e., the total number of ways you can order n different objects. To make our formulas consistent, we define $0! = 1$.

Example 2.5

Shuffle a deck of 52 cards. How many outcomes are possible? (In other words, how many different ways can you order 52 distinct cards? How many different permutations of 52 distinct cards exist?) The answer is $52!$.

Now, using the definition of $n!$, we can rewrite the formula for P_k^n as

$$P_k^n = \frac{n!}{(n-k)!}.$$

The number of k-permutations of n distinguishable objects is given by

$$P_k^n = \frac{n!}{(n-k)!}, \text{ for } 0 \leq k \leq n.$$

Note: There are several different common notations that are used to show the number of k-permutations of an n-element set including $P_{n,k}, P(n,k), nPk$, etc. In this book, we always use P_k^n.