Lecture 6

Andrei Antonenko

Feb 04, 2005

1 Counting

On the last lecture we computed the number of combinations, or number of all k-element subsets of the set with n elements. The formula obtained was

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}$$

Now we will count partitions.

1.1 Partitions

Assume there is a set A with n elements. Suppose that we want to break this set into r groups with given number of elements in each group: group 1 should have n_1 elements, group 2 should have n_2 elements, etc., and group r should have n_r elements. Furthermore, it should be true that

$$n_1 + n_2 + \cdots + n_r = 1.$$

In how many ways that can be done?

Since the 1st group should have n_1 elements, we can choose them in $\binom{n}{n_1}$ ways. Now we are left with $n - n_1$ elements, out of which we have to choose n_2. That can be done in $\binom{n-n_1}{n_2}$ ways. The same for other groups, and finally, for the r-th group we are left with $n - n_1 - \cdots - n_r$ elements, out of which we have to choose n_r. Thus, the total number of possible partitions is

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-n_2-\cdots-n_{r-1}}{n_r}.$$

We will simplify the formula above by using the expression for binomial coefficients:

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-n_2-\cdots-n_{r-1}}{n_r} =$$
$$= \frac{n!}{(n-n_1)!n_1!}\frac{(n-n_1)!}{(n-n_1-n_2)!n_2!}\frac{(n-n_1-n_2)!}{(n-n_1-n_2-n_3)!n_3!}\cdots\frac{(n-n_1-n_2-\cdots-n_{r-1})!}{(n-n_1-\cdots-n_r)!n_r!} =$$
$$\frac{n!}{n_1!n_2!\dots n_r!},$$

where the last equality is true because a lot of terms cancel out, and moreover, since $n = n_1 + \cdots + n_r$, $(n - n_1 - \cdots - n_r)! = 0! = 1$. So, the total number of partitions is

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}. \tag{1}$$

Now let's consider several examples.

Example 1.1. *How many different anagrams can be made out of word "TATTOO"? Anagram is a word obtained by the rearrangement of the letters of the initial word.*

We can solve this problem in 2 different ways.

First way is to consider all rearrangement of 6 letters of the word "TATTOO" – there are in total $6!$ *such rearrangements. Three letters "T" can be rearranged in* $3!$ *ways, and 2 letters "O" can be rearranged in* $2!$ *different ways. Thus, the total number of anagrams is equal to* $\frac{6!}{3!2!}$.

Another way to solve this problem is to use the formula for partitions. We have 6 slots available for letters. We need to decompose slots into three groups: one group for letter "T" of the size $n_1 = 3$*, second group for letter "O" of the size* $n_2 = 2$*, and the third group for letter "A" of the size 1. Thus, the total number of anagrams is equal to* $\binom{6}{3, 2, 1} = \frac{6!}{3!2!1!}$*, which is the same as in the previous method.*

Example 1.2. *Now we will solve the problem considered in one of the previous lectures using the formula for partitions. Assume there are 16 students in the class, 4 of them are boys, and 12 are girls. The professor divides students into 4 groups of 4 persons each. What is the probability that each group will have a boy?*

First of all, the total number of partitions here is $\binom{16}{4, 4, 4, 4} = \frac{16!}{4!4!4!4!}$.

Now, we will count the total number of partitions such that each group has one boy. This can be done in two stages.

First, let's arrange boys. First boy has 4 choices of groups, 2nd boy has 3 choices of groups, etc., thus the boys can be arranged by various groups in $4!$ *ways. Now let's arrange girls. Each of the 4 groups now has 3 empty slots, and we have total of 12 girls. Thus, we can arrange them in* $\binom{12}{3, 3, 3, 3} = \frac{12!}{3!3!3!3!}$.

Therefore, the total number of arrangements of students into 4 groups such that each group has a boy is equal to

$$\frac{4!12!}{3!3!3!3!}.$$

Thus, the probability of having an arrangement with boy in each group is equal to

$$\frac{\frac{4!12!}{3!3!3!3!}}{\frac{16!}{4!4!4!4!}} = \frac{4!12!4!4!4!4!}{3!3!3!3!16!} = \frac{4! \cdot 4 \cdot 4 \cdot 4 \cdot 4}{13 \cdot 14 \cdot 15 \cdot 16} = \frac{24 \cdot 4 \cdot 4}{13 \cdot 14 \cdot 15}$$

which is the same as the answer obtained before using conditional probability.

2 Discrete random variables

Most experiments in the real life are of numerical nature. For example, observing the temperature, the behavior of the stock, etc. are numerical – we measure some particular numbers. Some experiments however are not numerical. The outcome of the experiment, consisting of 5 coin tosses is not a number, but a sequence of Hs and Ts. Rolling two dice also does not give us a single number, but gives us a pair of numbers. Often it is useful to assign a single number to any outcome of the experiment. Given a particular experiment, **random variables** associate a particular number with each outcome of it. Mathematically speaking, **random variable** is a function from the set of possible outcomes of the experiment to the set of real numbers.

Example 2.1. *i). In the experiment consisting of 5 coin tosses, for each its outcome we can associate a number of heads, obtained during the experiment. In this case number of heads is a random variable, taking values 0, 1, 2, 3, 4, 5.*

ii). In the experiment of rolling a dice twice, we can consider a random variable which is equal to the maximum number rolled, or equal to the number of 6*s rolled, or equal to the sum of two numbers rolled.*

Random variable is called **discrete** if its range (set of values it can take) is finite, or countably infinite (like a set of integer or natural numbers).

Example 2.2. *Consider an experiment consisting of choosing a random number from the interval* $[-1, 1]$. *If we associate with the outcomes of this experiment the value of the number chosen, this random variable will not be discrete, since its range is the whole interval* $[-1, 1]$ *which is neither finite nor countable. But if with any chosen number* a *we associate the number* X *such that*

$$X = \begin{cases} 1, & \text{if } a > 0 \\ 0, & \text{if } a = 0 \\ -1, & \text{if } a < 0 \end{cases}$$

we will have a discrete random variable X.

3 Probability mass function

For each possible value of the discrete random variable we can compute the probability that the random variable is equal to it. But before making this statement precise, I will clarify the notation. By capital letters (X, Y, etc.) we will be denoting random variable. By small letters (x, y, etc.) we will be denoting the values of the random variables.

Definition 3.1. *The* ***probability mass function (PMF)*** *of the discrete random variable* X *is a function which assigns the probabilities to all possible values of the given random variable:*

$$p_X(x) = P(\{X = x\}). \tag{2}$$

In different words, for a given random variable X, *the value of the probability mass function* $p_X(x)$ *is equal to the probability of the event* $\{X = x\}$.

Let's consider examples.

Example 3.2. *Consider an experiment of tossing a fair coin twice. Let the random variable X be the total number of heads. Than,*

$$p_X(0) = P(\{X = 0\}) = P(\{TT\}) = \frac{1}{4};$$
$$p_X(1) = P(\{X = 1\}) = P(\{HT, TH\}) = \frac{1}{2};$$
$$p_X(2) = P(\{X = 2\}) = P(\{HH\}) = \frac{1}{4};$$
$$p_X(3) = P(\{X = 3\}) = P(\emptyset) = 0.$$

So, the PMF is the following:

$$p_X(x) = \begin{cases} 1/4, & \textit{if } x = 0, 2; \\ 1/2, & \textit{if } x = 1; \\ 0, & \textit{otherwise.} \end{cases}$$

Example 3.3. *Consider the experiment consisting of 2 rolls of a fair 4-sided die. Let X be a random variable, equal to the maximum of 2 rolls. The possible values for this random variable are $1, 2, 3, 4$. The PMF can be counted as:*

$$p_X(1) = P(\{X = 1\}) = P(\{11\}) = 1/36;$$
$$p_X(2) = P(\{X = 2\}) = P(\{12, 21, 22\}) = 3/36;$$
$$p_X(3) = P(\{X = 3\}) = P(\{13, 23, 33, 32, 31\}) = 5/36;$$
$$p_X(4) = P(\{X = 4\}) = P(\{14, 24, 34, 44, 43, 42, 41\}) = 7/36.$$

The property worth mentioning here is that the sum of all values of the PMF should be equal to 1: if the range of the values of the random variable X is A, then

$$\sum_{x \in A} p_X(x) = 1. \tag{3}$$

Now we will consider several examples of commonly used random variables.

4 Commonly used discrete random variables

4.1 Bernoulli random variable

Consider tossing a biased coin once. Assume that the probability of H is p, and the probability of the T is equal to $q = 1 - p$. Let's consider a random variable X which is equal to 1 in the case of H, and 0 in the case of T. This random variable has PMF

$$p_X(x) = \begin{cases} p, & \text{if } x = 1; \\ 1 - p, & \text{if } x = 0; \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

This easy random variable is called **Bernoulli random variable** with parameter p. This is an important r.v., and is used often when talking about the experiment with two outcomes (person sick or healthy, phone line is busy or occupied, person voted for or against George W. Bush). In many cases, instead of talking about heads and tails, people talk about "success" and a "failure" of the experiment, with p being a probability of success, and $q = 1 - p$ being a probability of failure.

4.2 Binomial random variable

Consider a sequence of a biased coin tosses of the length n. As above, assume that the probability of H is p, and the probability of T is $1-p$. For such a sequence, let's consider a random variable X equal to the number of Heads in the sequence. In this case, X can take values from 0 to n. The probability of having k heads in the sequence of n tosses was obtained at the end of the previous lecture, and is equal to $\binom{n}{k}p^k(1-p)^{n-k}$. Therefore, the PMF of X is

$$p_X(k) = \binom{n}{k}p^k(1-p)^{n-k}. \tag{5}$$

This random variable is called a **binomial random variable** with parameters p and n. From the property that the sum of all probabilities should be equal to 1, we have the following identity:

$$\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} = 1.$$

Again, here we can talk about a series of experiments, each resulting in either success or failure. The random variable is equal to the number of successes in the sequence of n experiments, where p is a probability of success, and $1-p$ is a probability of failure.

4.3 Geometric random variable

Consider an experiment in which we repeatedly toss a biased coin with probability of H being p until the first H. Let the random variable X be a number of toss on which the first Head is obtained. Probability that the first H will be obtained on the k-th toss can be computed by noting, that this events happens when we have $k-1$ Tails followed by Head. Since the probability of each T is $1-p$, and probability of H is p,

$$p_X(k) = (1-p)^{k-1}p, \quad k = 1, 2, \ldots. \tag{6}$$

This random variable is called **geometric** with parameter p. The sum of all values of PMF should be equal to 1. This can be easily verified by direct calculations:

$$\sum_{k=1}^{\infty}p_X(k) = \sum_{k=1}^{\infty}(1-p)^{k-1}p = p\sum_{k=0}^{\infty}(1-p)^k = p\frac{1}{1-(1-p)} = 1,$$

where we used a formula for the sum of geometric progression

$$\sum_{k=0}^{\infty}p^k = 1 + p + p^2 + \cdots = \frac{1}{1-p}. \tag{7}$$

In terms of successes, this random variable is equal to the number of the experiment, on which the first success is achieved.