LECTURE 5

§8. COMBINATORIAL ANALYSIS

Many of the basic concepts of probability theory, as well as a large number of important problems of applied probability theory, may be considered in the context of finite sample spaces. Therefore, many problems in probability theory require that we count the number of ways that a particular event can occur.

The mathematical theory of counting is formally known as *Combinatorial Analysis*. Let us briefly give the main notions of Combinatorial Analysis.

1. Basic Principle of Counting.

Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of m possible outcomes and if, for each outcome of experiment 1, there are n possible outcomes of experiment 2, then together there are $m \cdot n$ possible outcomes of the two experiments.

Proof: The basic principle can be proved by enumerating all the possible outcomes of the two experiments as follows:

$$\begin{array}{llll}
(1,1), & (1,2), & ..., & (1,n) \\
(2,1), & (2,2), & ..., & (2,n) \\
.... & .... & ..., & .... \\
(m,1), & (m,2), & ..., & (m,n)
\end{array}$$

where we say that the outcome is (i, j) if experiment 1 results in its ith possible outcome and experiment 2 then results in the jth of its possible outcomes. Hence, the set of possible outcomes consists of m rows, each row containing n elements, which proves the result.

2. Generalized Basic Principle of Counting.

If k experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes, and if for each of these n_1 possible outcomes there are n_2 possible outcomes of the second experiment, and if for each of the possible outcomes of the first two experiments there are n_3 possible outcomes of the third experiment, and if ...., then there are a total of $n_1 \cdot n_2 \cdot ... \cdot n_k$ possible outcomes of the k experiments.

3. Ordered Sequences or Permutations. Briefly, a permutation is an ordered arrangement of objects. In general, if k objects are chosen from a set of n distinct objects, any particular arrangement, or order, of these objects is called a *permutation*.
How many different permutations of k objects selected from a set of n distinct objects are possible? There are

$$(n)_k = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot (n-k+1)$$

different arrangements of n objects taken k.
This result can be proved by using the basic principle of counting, since the first object in the permutation can be any of the n objects, the second object in the permutation can then be chosen from the remaining $(n-1)$, the third object in the permutation is then chosen from the remaining $(n-2)$ etc. and the final object in the permutation is chosen from the remaining $n-k+1$. Thus there are $n \cdot (n-1) \cdot ... \cdot (n-k+1)$ possible permutations of n objects taken k at a time.
$(n)_k$ read, permutations of n taken k. In particular

$$(n)_n = n \cdot (n-1) \cdot ... \cdot 2 \cdot 1$$

different permutations of n objects taken n.
Since products of consecutive positive integers arise in many problems relating to permutations or other kinds of special selections, it will be convenient to introduce here the *factorial* notation, where

$$n! = n \cdot (n-1) \cdot ... \cdot 2 \cdot 1 \quad \text{(read, } n \text{ factorial).}$$

In particular, to make various formulae more generally applicable, we let $0! = 1$ by definition.
To express the formula for $(n)_k$ in terms of factorials, we multiply and divide by $(n-k)!$ getting

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

and

$$(n)_n = n!.$$

Example 14. All possible rearrangements of the numbers $A = \{1, 2, 3\}$ (or permutations of 3 taken 3) are

$$123 \qquad 132 \qquad 213 \qquad 231 \qquad 312 \qquad 321.$$

Suppose that from the set $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ we choose k elements and list them in order. How many ways can we do it? The answer depends on whether we are allowed to duplicate items in the list. If no duplication is allowed, we are sampling *without replacement*. If duplication is allowed, we are sampling *with replacement*. We can think of the problem as that of taking labeled balls from an urn. In the first type of sampling, we are not allowed to put a ball back before choosing the next one, but in the second, we are. In either case, when we are done choosing, we have a list of k balls ordered in the sequence in which they are drawn.
The generalized principle of counting can be used to count the number of different samples possible for a set of n elements. First, suppose that sampling is done with replacement. The first ball can be chosen in any of n ways, the second in any of n ways, etc., so that there are $n \times n \times ... \times n = n^k$ samples. Next, suppose that sampling is done without replacement. There are n choices for the first ball, $n-1$ choices for the second ball, $n-2$ for the third, ..., and $n-k+1$ for the kth. We have just proved the following lemma.

Lemma 3. For a set of size n and a sample of size k, there are n^k different ordered samples with replacement and $(n)_k$ different ordered samples without replacement.

Example 15. How many ways five children be lined up?

Solution: This corresponds to sampling without replacement. According to the formula for permutation, we obtain

$$(5)_5 = 5! = 120.$$

Example 16. Suppose that from ten children, five are to be chosen and lined up. How many different lines are possible?

Solution: From Lemma 3, there are

$$(10)_5 = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = 30,240 \qquad \text{different lines.}$$

Example 17. Suppose that a room contains n people. What is the probability that at least two of them have a common birthday?

Solution: Assume that every day of the year is equally likely to be a birthday, disregard leap years, and denote by A the event that there are at least two people with a common birthday. As is sometimes the case, it is easier to find $P(\overline{A})$ than to find $P(A)$. This is because A can happen in many ways, whereas $\overline{A}$ is much simpler. There are 365^n possible outcomes, and $\overline{A}$ can happen in $365 \times 364 \times ... \times (365 - n + 1)$ ways. Thus,

$$P(\overline{A}) = \frac{365 \times 364 \times ... \times (365 - n + 1)}{365^n}$$

and

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{365 \times 364 \times ... \times (365 - n + 1)}{365^n}.$$

The following table exhibits the latter probabilities for various values of n:

Table

n	4	16	23	32	40	56
P(A)	0.016	0.284	0.507	0.753	0.891	0.988

From the Table, we see that if there are only 23 people, the probability of at least one match exceeds 0.5.

We now shift our attention from counting permutations to counting combinations. Here we are no longer interested in ordered samples.

4. Combinations. Consider n element set. How many possible subsets of size k ($k \leq n$) can be formed?

The number of subsets of size k that may be formed from the members of a set of size n is

$$\binom{n}{k} = \frac{n!}{(n-k)!\ k!}, \qquad \text{(read, } n \text{ choose } k\text{)} \tag{15}$$

Therefore, $\binom{n}{k}$ represents the number of possible combinations of n objects taken k at a time, or the number of different groups of size k that could be selected from a set of n objects when the order of selection is not considered relevant.

Remark 4. In taking samples sequentially from a discrete sample space, the sampling may be done either *with replacements* or *without replacements*; that is when an element is drawn from the set it is either returned or not returned to the set before another element is drawn. In a set of n elements, the number of ordered samples of size k then is n^k for sampling with replacements, whereas in sampling without replacements, the corresponding number of ordered samples is $(n)_k$.

Now we ask the following question: If k objects are taken from the set of n objects without replacement and disregarding order, how many different samples are possible? From the general principle of counting, the number of ordered samples equals the number of unordered samples multiplied by the number of ways to order each sample. Since the number of ordered samples is $(n)_k$ and since a sample of size k can be ordered in $k!$ ways, the number of unordered samples is (15) and therefore, coincides with $\binom{n}{k}$.

5. Unordered samples with replacements. Suppose that a sample of size k is to be chosen from a set with n elements. In sampling without replacement, no elements may be chosen more than once, so that the k items in the sample will all be different. In sampling with replacement, a member may be chosen more than once, so that not all of the k items in the sample need be different. Indeed it is possible that the same item might be chosen every time, in which case the sample would consist of a single item repeated k times.

Lemma 4. The number of ways to form (unordered) combinations of length k from a set of

n *distinct objects, replacements allowed, is equal to*

$$\binom{n+k-1}{k}$$

Proof: Each k–combination of a set with n elements when repetition is allowed can be represented by a list of $n-1$ bars and k stars. The $n-1$ bars are used to mark off n different cells, with the ith cell containing a star for each time the ith element of the set occurs in the combination. For example, a 6–combination of a set with four elements is represented with three bars and six stars. Here

$$* * \mid * \mid\mid * * *$$

represents the combination containing exactly two of the first element, one of the second element, none of the third element, and three of the fourth element of the set.
As we have seen, each different list containing $n-1$ bars and k stars corresponds to an k–combination of the set with n elements, when repetition is allowed. The number of such lists is $\binom{n+k-1}{k}$, since each list corresponds to a choice of the k positions to place the k stars from the $n+k-1$ positions that contain k stars and $n-1$ bars.

Let us consider examples.

Example 18. How many people must you ask in order to have a $50:50$ chance of finding someone who shares your birthday?

Solution: Suppose that you ask n people. Let A denote the event that someone's birthday is the same as yours. Again, it is easier to work with $\overline{A}$. The total number of outcomes is 365^n, and the number of ways that $\overline{A}$ can happen is 364^n ways. Thus,

$$P(\overline{A}) = \frac{364^n}{365^n}$$

and

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{364^n}{365^n}.$$

In order for the latter probability to be 0.5, n should be 253.